

Content

| | |
|---|----|
| Preface..... | 7 |
| Executive Summary..... | 9 |
| 1.Introduction..... | 13 |
| 2. The working procedure..... | 15 |
| 3. Overview of the comments to the MERAG fact sheets..... | 17 |
| 3.1 General comments brought up by several reviewers at the Oslo workshop..... | 17 |
| 3.1 The fact sheets and related chapters in the background document. | 18 |
| ANNEX 1:..... | 27 |
| Fact sheet 1 - “Risk characterization – general aspects“..... | 27 |
| Aim and structure of the fact sheet..... | 27 |
| Comments..... | 27 |
| ANNEX 2:..... | 29 |
| Fact sheet 2- “Exposure assessment”..... | 29 |
| Comments..... | 29 |
| Conclusions..... | 33 |
| References..... | 35 |
| ANNEX 3:..... | 37 |
| Fact sheet 3“Data compilation, selection and derivation of PNEC values for the risk assessment of different environmental compartments (water, sewage treatment plants, soil, sediment)” and Chapter 1 in the background document: “Metals are intrinsic parts of the environment”..... | 37 |
| Aim with the fact sheet..... | 37 |
| Comments and suggestion..... | 37 |
| References..... | 46 |
| ANNEX 4:..... | 49 |
| Fact Sheet 4 - “Marine risk assessment” and chapter 4 of Background Document “Biodiversity and sensitivity of marine versus freshwater systems”..... | 49 |
| 1. Introduction, aims and structure of the fact sheet..... | 49 |
| 2. Comments to fact sheet 4: Marine Risk Assessment..... | 50 |
| 3. Comments to “One-pager” 13: Use of Freshwater Ecotoxicity Data to Determine a Marine PNEC..... | 52 |
| 4. Comments to “Metal Risk Assessment Guidance Document: Background Document (section 4: Biodiversity and sensitivity of marine vs. freshwater systems)”..... | 53 |
| 5. Conclusions..... | 53 |
| 6. References..... | 53 |
| ANNEX 5:..... | 55 |
| Fact sheet 5 - “Incorporation of bioavailability for water, (soils) and sediments” and background Chapter 2 in the background document: “Bioavailability of metals/metal compounds”..... | 55 |
| Aim with the factsheet..... | 55 |
| Incorporation of bioavailability in fresh-water..... | 55 |
| - Short description on the approach and proposed methods..... | 55 |
| - Comments and suggestion..... | 56 |
| Incorporation of bioavailability of metals in sediments..... | 60 |
| - Short description of the approach and proposed methods..... | 60 |
| - Comments, questions and suggestions to the fact sheet..... | 60 |

| | |
|--|----|
| - Comments to Background document ch. 2.3 | 63 |
| References | 64 |
| ANNEX 6: | 67 |
| Fact Sheet 6 - "Incorporation of bioavailability for soils" and Chapter 2.4 in the background document "Bioavailability of metals/metals compounds in the terrestrial compartment" | 67 |
| Aim of the fact sheet | 67 |
| Short description on the approach and proposed methods | 67 |
| Comments and suggestions | 67 |
| Conclusion | 72 |
| References | 72 |
| ANNEX 7: | 73 |
| Fact sheet 7: "Uncertainty analysis", and Chapter 6 on the Background document: "Examples of probabilistic techniques and quantification of uncertainty in metal risk assessments" | 73 |
| Aim with the Fact sheet | 73 |
| Comments and suggestions | 73 |
| Conclusions | 74 |
| Comments to Chapter 6 in the Background document (<i>Examples of probabilistic techniques and quantification of uncertainty in metal risk assessments</i>) | 75 |
| Reference | 76 |
| ANNEX 8: | 77 |
| Fact sheet 8: "Data compilation, selection and derivation of ecotoxicity reference values. MERAG Program- Building Block: Classification for effects on the aquatic environment." (Version June, 2006) | 77 |
| Aim with the fact sheet | 77 |
| Short description on the approach and proposed methods | 77 |
| Comments and suggestions | 77 |
| Conclusion | 78 |
| Reference | 79 |
| Sammanfattning | 81 |
| Abbreviations and acronyms | 85 |

Preface

This work was initiated by the Nordic Risk Assessment Project (NORAP), which got founding from the Nordic Chemicals group, a subgroup to the Nordic Council of Ministers.

The present NORAP work is mainly directed at following and contributing to the development of Technical Guidance Documents within the REACH Implementation Projects, RIPs (e.g. projects on: Information Requirements and Preparing the Chemical Safety Report). During 2006, NORAP's work has among other been focused on reviewing the metal risk assessment guidance documents for environment and health, MERAG and HERAG respectively, developed by the metals industries in co-operation with the government in United Kingdom. The present report constitutes the result of the reviewing of the environmental risk assessment guidance documents.

The reviewing has been performed by: Torsten Källqvist, Ketil Hyl-land, Jens Skei and Tuomo Saloranta from the Norwegian Institute for Water Research (NIVA), Hans Borg and Göran Lithner from Department of Applied Environmental Science (ITM) in Sweden, Janeck J. Scott-Fordsmand from the National Environmental Research Institute (DMU) in Denmark, and the Nordic Classification and Labelling Group (led by Jonas Falck). The individual reports from the different reviewers are included in this report as Annex 1-8. The main conclusions from these report, as well as the working procedure, have been summarised by Helena Parkman, Swedish Chemicals Agency (KemI) with input from the NORAP environmental group, including Ivar Lundberg, KemI, Toralf Kaland, Norwegian pollution Control Authority (SFT), Jaana Heiskanen, Finnish Environment Institute (SYKE) and Henrik Tyle, Danish environmental Protection Agency (MST).

Executive Summary

The aim of the present NORAP project was to critically review proposals on principles for risk assessment of metals (MERAG), developed by the metals industry, in co-operation with the UK Government during 2005-2006. This review was considered urgent, in order to be prepared for a possible proposal to implement MERAG within the ESR (Existing Substances Regulation, EEC 793/93) as well as in the Technical Guidance Document on Chemical Safety Assessment, which will be developed for use under the REACH regulation.

The MERAG project is structured around 5 main building blocks (generic concepts, classification and ranking, effects assessment, exposure assessment and risk characterisation) constituted of fact sheets covering different key aspects of risk assessment methodology. A background document provides motivation, justification and explanation of the proposals made.

The individual fact sheets have been reviewed by different Nordic consultancies, and their evaluations are added as stand-alone documents in Annexes 1-8 of the report. The NORAP steering group has summarised the main conclusions from the work, but these do not necessarily reflect the formal position of the authorities of the Nordic countries.

The project has not only resulted in the present report, but has also increased the involvement and understanding among Nordic researchers, of the environmental risk assessment procedure applied within certain EU regulations.

It is recognised that metals need specific guidance not only because they cover a distinct group, inorganics, but also for reasons such as natural backgrounds, essentiality, and long historical use. For some metals it might be necessary to take bioavailability into account in risk assessments. The MERAG documents constitute good starting points for further development of guidance.

A workshop was held in Oslo (February 2006) with all the consultancies. One common conclusion from this gathering was that Nordic conditions are not sufficiently covered in the different bioavailability correction models. For the metals where bioavailability corrections have been carried out until now (Cu, Zn, Ni, and partly Cd and Pb), biased databases on abiotic conditions have been used to set “typical” and “worst case” scenarios as well as the limits for bioavailability models (CEC/pH, BLM and AVS). Lately the FOREGS data have been made available and may be used to derive typical values for abiotic factors in European freshwaters. The metal concentrations in this database may, with careful considerations, be used as surrogates for current background levels.

It was also concluded at the workshop that revisions and improvement of the scientific quality is needed in the MERAG background document.

A third conclusion was that the MERAG documents are imbalanced in terms of criteria for data selection. Much more effort have been put on the discussion of criteria for the selection of acceptable toxicity data, compared to the criteria for acceptance of data for exposure analysis. However, the NORAP group recognises that this is not a metal specific issue.

The general concept of a tiered approach in the risk characterisation strategy, fact sheet No1, is supported as long as the first tier is based on conservative estimates of PEC and PNEC. The need to use an assessment factor not lower than 2 (due to remaining uncertainty) in the derivation of PNEC from SSD, is stressed. The concept of using bioavailable fractions in estimates of PEC and PNEC will improve the basis of risk assessments. However, the limitations of the models used (e.g. BLM) must be recognised. The suggestion to refine the risk assessment by using a probabilistic approach is in general supported.

In the fact sheet No 2 on exposure the reviewing was focused on exposure assessment, using measured data. The reviewer concludes that guidance on criteria for quality and representativity of analytical data is needed. Various terms for 'background' concentrations should be more strictly defined. A discussion on the importance of seasonal variation should be included in the guidelines for data selection. When deriving the worst case scenario the 90th percentile is used, however, it would be more logic to use the 95th percentile, since the data is compared with the 5th percentile of the effects data.

In the review of fact sheet 3 on the effects assessment it is concluded that the issue on acclimatization/adaption suffers from large information gaps. There is e.g. a need for information on actual ranges of metal background concentrations, as well as factual OCEE curves for organisms. This is especially important in relation to the introduction of the metalloregion concept, which also should be explained in much more detail. When PNEC is derived from an SSD the reviewer stresses that an assessment factor >2 should be applied, since much uncertainty is not covered by the SSD. Examples of such uncertainties are: additional exposure routes, uncertainties of used speciation models, and importance of keystone species and ecosystem processes. Finally more justification than what is presented in the fact sheet is needed for lowering the number of taxonomic groups and NOECs (compared to what is stipulated for the water compartment), when applying SSD for the sediment and soil compartment.

The reviewer of the fact sheet No 4, on marine risk assessment, is very critical to the description of the marine system and the proposals in the fact sheet, but also to the generic method as explained in the TGD. It is unlikely that toxicity testing with only freshwater species will be protective of all marine species, and it is not possible to provide an appropriate

assessment factor for extrapolating freshwater data to the marine environment.

Fact sheet No 5 deals with bioavailability in water and sediment. The BLM concept based on chemical modelling for transformation of soluble metals to bioavailable species in water is good, pioneering and builds on new research. However, there is a tendency in the MERAG documents to underestimate the uncertainties of the BLM modelling, e.g. the extrapolation from a few test species for which BLMs exist to other test species and to the whole ecosystem. This should be discussed much more.

There is an overall agreement that concentrations of total metals in sediments are possibly only indirectly related to environmental risk. The guidance on bioavailability correction (fact sheet 5) is focused on sulphide binding metals and the SEM/AVS concept. However, this approach is marred by great uncertainty, e.g. regarding the relevance of the method for benthic species with different burrowing behaviour, and the dependence on spatial and temporal variation and stability of AVS in the sediments of the recipient under investigation. The SEM/AVS work has increased our understanding of some processes in sediments, but should not be presented as the only available risk assessment tool. Alternative methods might be to relate the toxicity to the metal content in overlaying water, or to make bioavailability corrections, similar to what is proposed for water and soil.

Fact sheet No 6 deals with how to implement bioavailability into the ecological risk assessment of metals in soil. In general the suggested approach will enable a more refined risk assessment than previously possible, as soil-type models are introduced. However, the approach is solely bioavailability oriented and ignores other important issues, such as mixture toxicity, competition and more ecological and climatic aspects. A crucial point is for which species the models are derived and whether the models can be applied to other species. MERAG proposes to apply a *lab-to-field factor* to the toxicity data. However, the suggested method to derive this extrapolation factor does not cover all aspects of lab-to-field extrapolation/differences. There is a tendency to only include factors that may reduce the risk.

The fact sheet No 7 on uncertainty analysis, introduces quite well most of the cornerstones of uncertainty analysis. However, more rigorous definition of some key terms is needed. More emphasis should be put on the fact that some types of uncertainty are very difficult to estimate quantitatively, and/or are irreducible, and that therefore all uncertainty analyses will basically remain more or less uncertain.

The reviewers of fact sheet 8 on classification concluded that there is no need for development of new guidance for environmental classification of metals and metal compounds, since globally agreed guidance already exist. Hence, the fact sheet No 8, on classification, should be replaced by reference to agreed guidance.

1. Introduction

The MERAG (Metals Environmental Risk Assessment Guidance) program was initiated by the metals industry represented by Eurometaux (European non-ferrous metals industry) and ICMM (International Council for Mining and Metals), and aims to “provide the regulatory community at regional and international level with scientific and regulatory guidance on the most advanced status of environmental Risk Assessment concepts for metals”. The program achieved managerial and review sponsorship by UK-DEFRA (i.e. the environment ministry in United Kingdom). Regulatory representatives from several other countries including Netherlands, USA, Canada and Italy provided additional technical guidance and review. The outcome of the project was discussed at an open workshop in London (May 2005), with participants from EU, USA and Canada representing industry, regulatory bodies and independent scientists.

The MERAG project is structured around 5 main building blocks (generic concepts, classification and ranking, effects assessment, exposure assessment and risk characterisation) constituted of 9 *fact sheets* covering different key aspects of risk assessment methodology for metals requiring specific guidance. The fact sheets summarise the consensus recommendation of the consultant (EURAS) the industry and external review panel, while a *background document* provides motivation, justification and explanation of the proposals made.

At the London workshop fact sheet No 8 on ‘Secondary poisoning’ was considered ‘Not mature enough’ and was hence not recommended as a guidance document. The other documents were considered useful as guidance documents after minor or more substantial revisions and/or clarifications. It is the intention of the sponsors of the program to introduce the work products within EU and OECD chemicals management programs for consideration.

Although the final versions of the fact sheet were not available yet, the present Nordic project was initiated with the aim to critically review the MERAG proposal. This was considered urgent, in order to be prepared for a possible proposal to implement MERAG within the ESR (Existing Substances Regulation, EEC 793/93) as well as in the Technical Guidance Document on Chemical Safety Assessment, which will be developed for use under the REACH regulation. In addition, “horizontal metal issues” with more or less connection to the MERAG fact sheets has been and will be discussed at technical meetings within the ESR (Existing Substances Regulation, EEC 793/93) framework during 2006 and 2007. The MERAG fact sheets have already, and will probably more intensively in the near future, be put forward by the industry in the preparatory

work for REACH, the so-called REACH Implementation Projects (RIPs). The opinions in the present report will be used to feed into the discussions within this work.

2. The working procedure

The individual fact sheets have been reviewed by different Nordic consultancies. The reviewers presented their draft opinions at a workshop in Oslo at end of February 2006, which also was attended by some external experts and delegates from the NORAP steering group. The final evaluation reports of the reviewers, including the discussions from the Oslo-workshop, are added to this summary report as independent annexes (Annex 1-8). These annexes are stand-alone documents. Based on the annexes the NORAP steering group has made an overall summary in Chapter 3 of this report. Hence, the standpoints in the document do not necessarily reflect the formal position of the authorities of the Nordic countries.

The development of guidance for risk assessment of metals is a living continuous process, and since the contracted consultancies made their reviews of the fact sheets, the fact sheets have been revised and two new versions have been made available, the first dated July 2006, and the second September 2006. Not many revisions were made between the June version and the September version, but the numbering of the fact sheets was changed. The reviewers of fact sheet 1 (on risk characterisation) and fact sheet 8 (on environmental classification) have commented on the versions from June 2006. For the other fact sheets the comments were made on the versions from May 2005, but most of the comments are still valid for the new versions.

In the meantime, also so-called “one-pagers” with horizontal metals issues, based partly on the MERAG documents, have been produced by the metals industry and various metals organisations, as background documents for discussions held at the Technical Committee for New and Existing substances. Comments to some of these “one-pagers” have been included in the reports from the reviewers (Annex 1-8).

This process has not only resulted in the present report with comments from independent researchers on the developing principles for environmental risk assessment of metals. Another gain has been the involvement of Nordic researchers and their increased understanding for the environmental risk assessment of metals within the EU, bridging the gap between research and the exercise of authorities

3. Overview of the comments to the MERAG fact sheets.¹

Several years of experience with assessing metals within the Existing Substances Regulation is the platform for the development of the MERAG. The existing “Technical Guidance Document on Risk Assessment” for new and existing substances is developed mainly for organic chemicals. It is recognised that metals need specific guidance because they cover a distinct group (inorganics), but also for several other reasons such as natural backgrounds, essentiality, and long historical use. For some metals it might be necessary to take bioavailability into account in risk assessments, however, the methods used shall not result in larger uncertainty compared to the standard approach. It is therefore valuable with general, well-reasoned and validated guidelines.

Much effort has been put into the production of the MERAG documents and they constitute good starting points for further development of guidance. In this review we have focussed on issues where further discussions, clarification and or scientific evidence are needed.

3.1 General comments brought up by several reviewers at the Oslo workshop

A general comment on the “background document” is that the text lacks depth and explanations and many examples are presented in a very brief manner. The text is also generally rather incomplete and draft-like (missing references, clutter in figure numbering & labels). In addition references cited in the text are often not from peer-reviewed journals or could not even be found in the reference list. Thus major revisions and improvement of the scientific quality should be made to this document before it could be of any larger value to a wider audience.

Another general and important comment from the Oslo workshop was that Nordic conditions are not sufficiently covered in the different bioavailability correction models, and it is recommended that research is initiated to validate the different models for “Nordic conditions”.

Finally, it was concluded that the MERAG documents are imbalanced in terms of criteria for data selection. Much more effort have been put on the discussion of criteria for the selection of acceptable toxicity data,

¹ The reviewers comments do not necessarily reflect the formal position of the authorities of the Nordic countries.

compared to the criteria for acceptance of data for exposure analysis. This regards monitoring data as well as measurements of abiotic parameters in toxicity tests, and is especially important for metals for which bioavailability corrections are made

3.1 The fact sheets and related chapters in the background document.

Fact sheet No 1 on risk characterisation

This fact sheet integrates the principles outlined in detail in other fact sheets into a risk assessment strategy. A tiered approach is proposed where the risk assessment can be refined along two different lines. The first line starts with assessment based on total concentrations, where the refinements include considerations of dissolved fractions, bioavailable fractions and, finally, the metalloregion concept. The second line of refinement goes from deterministic to probabilistic risk assessment, with field validations as a further refinement.

The general concept of a tiered approach is supported as long as the first tier is based on conservative estimates of PEC and PNEC. This implies that further refinement of the risk assessment will not be required when a risk quotient less than one has been obtained.

The concept of using bioavailable fractions in estimates of PEC and PNEC will improve the basis of risk assessments. However, the limitations of the models used (e.g. BLM) must be recognised.

The purpose of the final refinements step, the metalloregion approach, is that geographical/geological as well as biological characteristics shall be considered in risk assessments on a regional or local scale. The rationale for this approach is that the biota may be adapted to regional or local variations in natural background concentrations. However, the acclimatization/adaptation concept suffers from large information gaps. In fact, the ultimate approach to adopt the metalloregion concept would be to produce a complete set of data on chronic toxicity to “endemic” species. This would require a large number of tests to be able to perform ‘endemic’ SSD analysis for each metalloregion.

The suggestion to refine the risk assessment by using a probabilistic approach is in general supported. However, it should be recognised that an analysis based on the exposure concentration distribution (ECD) and species sensitivity distribution (SSD) may not give a complete picture of the ecological consequences as long as only the number of species affected, and not which species that are affected is considered. For instance, effects on “keystone” species or species of particular importance from an ecological point of view must not be ignored. In addition, analysis based on ECD will not allow assessment of effects due to peak concentrations,

which may be fatal to organisms during specific parts of the life cycle. Variations in abiotic factors should also be included in the analysis.

Finally, the reviewer stresses that there is no general acceptance of the direct use of the 5%-ile from an SSD as PNEC. Instead, according to TGD, an assessment factor (of 5-1) should be used for derivation of a PNEC from an SSD. This concept must also be included in the probabilistic risk assessment. Since certain effects of ecological significance, such as reduced competitive ability or changes in behaviour will not be revealed by single species laboratory testing, there will always be a remaining uncertainty in derivation of PNEC from an SSD. For this reason an assessment factor of 1 can not be justified and a lower limit of 2 as the assessment factor when PNEC is derived from SSD is recommended.

Fact sheet No 2 on exposure assessment

This fact sheet deals with the exposure assessment, based on modelled and measured data. However the reviewer has focused his comments on the part dealing with exposure assessment using measured data (section 3). In this section guidance is given on how to select and handle monitoring data for derivation of PEC, however, acceptance criteria for the data used for exposure analysis and background levels of metals are largely missing. In addition, the term “natural background” is used alternately with “ambient concentrations” and “background concentrations” in the documents. A more strict definition of “natural” should be presented. The data that are referred to as natural background levels of metals, in the fact sheet should not be used for that purpose, since they are in general extremely high compared to e.g. levels in the FOREGS database.

In exposure assessments for the water compartment, MERAG recommends to use the concentrations of dissolved metals, which in the fact sheet is defined as the fraction passing a 0.45 µm filter. However, such a filtrate does not represent a truly dissolved fraction of metals since it includes also fine particulate colloidal matter.

The seasonal variation of metal concentrations in especially running waters could be considerable. The influence of seasonal variations on sampling programs is only briefly discussed. More detailed instructions for annual sampling frequency and sampling seasons should be included in the guidelines for the determinations of ECD, PEC and in the criteria for data selection.

The reviewer also comments that the use of 90th percentiles for both effluent and receiving waters may exclude elevated (maximum) values, which could be of ecotoxicological significance in a recipient. It would be more logical to use the 95th percentile for the worst case scenarios. That would also be more consequent when comparing with the statistical treatment of data suggested for the PNEC derivation in the effects as-

assessments (HC₅). This is however not a metals specific issue, since it is proposed in the TGD to use the 90th percentile.

Finally, the reviewer gives several proposals on how to improve the Questionnaire for exposure analysis (Annex 2).

Fact sheet No 3 (and chapter 1 of Background document) on effects assessment

The aim of the fact sheet No 3 is to give guidance in how to account for metal specific considerations in the derivation of PNEC.

Measured test concentrations are considered more reliable than nominal concentrations. However, while much attention is paid to secure the quality of toxicity data, the quality of measured chemical data in the toxicity tests is not discussed. Criteria for quality and representativity of analytical data need to be elaborated.

The acclimatization/adaptation concept suffers from large information gaps. The shortage of information may hamper the implementation of principles/ methods suggested in the fact sheet, e.g. that: “The background concentrations in culture media ideally should be representative for the organism and area under observation”, or “The essential metal concentration in culture medium should be at least equal to the minimal concentration causing deficiency for the test species”. There is need for objective information showing the factual ranges of metal background concentrations in selected areas, and factual OCEE (Optimal Concentration for Essential Elements) curves for selected species of aquatic organisms. Metal deficiency (and toxicity) in aquatic organisms need to be concretised and illustrated by importance of macronutrients and other modifying factors. In the background document (Chapter 1) there is a strong emphasis on essentiality. This may obscure the fact that many metals are not essential, and that different organisms may have different requirements.

The acclimatization/adaptation issue is related to the metalloregion concept, presented in the background document. This concept has to be explained much more in detail, i.e. substantiated in relation to environmental risk assessment.

The reviewer has pinpointed some issues regarding the ecological relevancy of tests, test species and test systems. One very important, metal specific problem is that metal toxicity in algae may be underestimated if nutrient rich test media is used (according to the guideline).

Other recommendations, for the effects assessment of the soil compartment, are that secondary consumers (e.g. mites) should be added as an additional taxonomic group to assess, and that enzymatic parameters should not be disregarded as effects parameter.

The reviewer proposes that any sensitive keystone species and ecosystem processes should be considered. They could in first hand be consid-

ered by increasing their weight in the derivation of PNEC-values (when considering the assessment factor). Results from mesocosm studies should be considered in the final risk assessment procedure, if they cover a relevant ecosystem and concentration interval. Regarding endpoints, any effects on individual species could be used from mesocosm studies in addition to total biomass, species diversity, and species richness. Special attention should be paid to *functional parameters*.

The reviewer also proposes to take ecosystem processes into account, in the choice of assessment factor. Another proposal is to compensate for large collective risk with a higher assessment factor, the “scale factor”, which should be applied to large water-bodies, e.g. sea gulfs, great lakes and open sea.

When calculating the PNEC for data rich substances the SSD (species sensitivity distribution) can be used according to TGD. An assessment factor between 5 and 1 should be applied to the derived HC₅. The reviewer considers that the factor should be >2 since much uncertainty is not covered by the SSD. Examples of such uncertainties are additional exposure routes, uncertainties of used speciation models and the factors (keystone species, ecosystem processes, and scaling factor) mentioned above.

Regarding the use of SSD for the sediment and soil compartment, MERAG proposes that the SSD should cover a minimum of three taxonomic groups and only four NOEC values for different species. For the aquatic compartment at least 10 NOECs for different species covering at least 8 taxonomic groups is needed for using SSD method. Therefore more justification is needed for lowering the number of taxonomic groups and NOECs when applying SSD for the sediment and soil compartment.

Fact sheet No 4 on marine assessment.

The aims of the fact sheet are to characterise marine ecosystems (abiotic and biological components), compare species diversity in freshwater and marine ecosystems, compare species sensitivity between the two ecosystems and to develop a proposal for PNEC derivation for marine ecosystems when there is a paucity of data.

It is acknowledged that toxicity data for the marine environment is scarce, and there is guidance in TGD on how to use freshwater toxicity data for PNEC derivation for the marine environment. However, concern has been raised regarding the relevance of this guidance.

The reviewer of this fact sheet is very critical to the proposals in the fact sheet, but also actually to the generic method as laid out in the TGD.

The reviewer questions the TGD approach to use freshwater results only, for deriving the marine PNEC. It is unlikely that toxicity testing with only freshwater species will be protective of all marine phyla. The arguments are given in Annex 4. One example is that freshwater and ma-

rine organisms differ in physiology, the most important difference being the difference in ion regulation and hence potential ion pumps and other potential metal transporters in gills or skin.

The reviewer advocate that only marine data should be used for marine risk assessment, and consider it not possible to provide an appropriate extrapolation factor should freshwater data be considered. However, if this approach is still used, the factor should presumably not be below the factor (10) suggested in the TGD to be protective for marine species. The argumentation for this pertains to the greater diversity of physiology in marine species, and will have a general application for both organic and inorganic contaminants.

In addition the reviewer considers the description of the marine chemistry of trace metals in the fact sheets as scientifically weak, e.g. does not correctly reflect the existing knowledge about metal bioavailability in marine ecosystems. The discussion of species diversity and taxonomic variation in freshwater compared to marine ecosystems should be improved, and should be based on more recent literature.

Fact sheet No 5 (and chapter 2 of Background document) on incorporation of bioavailability for water and sediments.

The aim with the fact sheet No 5 is to give guidance in methods for taking into account bioavailability of metals in the aquatic compartment, in order to “reduce the uncertainty and to increase the ecological relevance”.

It is stated in the fact sheet that the datasets of abiotic factors and environmental concentrations, used for bioavailability correction, should be “representative for and should cover the specific area under investigation”. However representativity and coverage are not explicitly explained in the text. For the metals where bioavailability corrections have been carried out until now (Cu, Zn, Ni, and partly Cd and Pb), biased databases have been used to set “typical” and “worst case” scenarios as well as the limits for bioavailability models (BLM and AVS). For instance, in the voluntary RAR for copper, data from only six EU-countries were used. Instead data from randomised sampling and high spatial coverage should be used to increase the realism of such scenarios. The FOREGS data base may be used to choose typical values for abiotic factors and the metal concentrations in this database may be used as surrogates for current background levels in remote areas. Surrogates for natural background levels may be found in far north. However, the definition of current and natural background level is a complicated issue.

MERAG proposes a stepwise refinement approach to take into account bioavailability in the water compartment. First, transform total concentrations to dissolved concentrations. If risk is indicated refine further by estimating the free ion metal fraction by physical-chemical speciation models. A final refinement step is to make use of toxicity based

models. The toxicity based model that has been put forward within MERAG, is the BLM model. This concept based on chemical modelling for transformation of soluble metals to bioavailable species in water is good, pioneering and builds on new research. However, BLM models only exist for a few metals and do not cover all abiotic conditions. In addition, there is a tendency in the MERAG documents to underestimate the uncertainties of the modelling (Annex 5). For instance problems in estimating complexation of cat ions like Al and Fe to DOC, is not taken into account, which may result in underestimation of free metal ions. It has also recently been demonstrated that the speciation model (WHAM) can predict the distribution of the dominant species of the metal fairly well, but when the free ion represent a very small fraction of the total concentration, it may be underestimated by several orders of magnitude (see Annex 5). Hence, there is an urgent need to validate the speciation models as well as the BLMs by measuring different metal species and toxicity in different types of natural water.

Finally, large uncertainty remains regarding the extrapolation from a handful of test species for which BLM models exist, to other test species, and to endemic species and ecosystems to be protected.

There is an overall agreement that concentrations of total metals in sediments are not directly related to environmental risk. Quantitative assessment of bioavailability of metals is important in risk assessment of contaminated sediments. However, prediction of the bioavailability of metals in sediments is far from straight forward.

MERAG differentiate between sulphide binding and non-sulphide binding metals. The guidance is focused on sulphide binding metals and the SEM/AVS concept. For the non-sulphide binding metals, it is proposed to explore if a relationship can be established between the observed toxicity levels and the presence of organic carbon, or with other sediment ligands such as Fe/Mn oxyhydroxides. However, this should not exclusively be an issue for elements which do not bind to sulphides.

The SEM/AVS approach is marred by great uncertainty for several reasons that are listed in Annex 5. One example is that it appears difficult to decide what sampling procedure would be most relevant to assess bioavailability and toxicity to organisms living close to sediment water interfaces, as well as for the burrowing fauna that often oxidize their burrows. The AVS concentrations used should be representative for where the benthic fauna thrive, but the sampling procedure is of uttermost importance, since almost any value for AVS is possible to obtain by adjusting the sampling depth/layer. Another example is that bioaccumulation of metals have been demonstrated from sediments where the metal should not have been bioavailable according to SEM/AVS. The reviewer also states that the solubility products used for metal sulphides in the proposed SEM/AVS approach for ranking metals, assume pure metal sulphide phases, which may not be entirely relevant in natural aquatic systems.

If the SEM/AVS method is to be used, the relevance for species with different burrowing behaviour need to be discussed, as well as the impact of spatial and temporal variations of AVS within a recipient. Therefore it is very questionable to apply a generic (worst case) AVS concentration for all recipients. In addition, in the risk characterisation an AVS normalized PEC must be compared with an AVS normalised PNEC.

Hence, the SEM/AVS work has increased our understanding of some processes in sediments, but other risk assessment tool should also be considered. It has e.g. recently been shown (in the risk assessment report on nickel) that the SEM/AVS approach does not work for benthos inhabiting the sediment-water interface, but toxicity relates better to the overlaying water content. Another alternative would be to make bioavailability corrections, similar to the approaches proposed for water and soil (e.g. BLM).

Fact sheet No 6 on incorporation of bioavailability for soils

For implementing bioavailability into the ecological risk assessment of metals in soil, three main issues are introduced: 1) a translation of toxicity from total to pore-water or free-ion concentrations, 2) a soil-type-correction of toxicity data and 3) a lab-to-field correction factor.

In general the suggested approach will enable a more refined risk assessment than previously possible, as soil-type models are introduced. However, the approach is solely bioavailability orientated and ignores other important issues, such as mixture toxicity, physical stress factors (e.g. climate), competition and more ecological aspects. A discussion of pros and cons (see Annex 6) of the applied approaches should be included, at least in the background document.

For *converting total concentrations* to pore-water or free-ion concentrations, different speciation models are proposed. The approach is based on more or less mechanistic models. It assumes a specific exposure route and that a point estimate of exposure reflects long-term toxicity. This has not been discussed and needs to be verified for a wide range of organisms. In addition, when more complex models are introduced there is a loss of clarity. The magnitude of the cumulative uncertainty should be clearly shown and the model should be validated on an external data-set.

A *soil-type normalisation* approach is suggested on the PNEC side by using regression models for the soil-parameters, e.g. toxicity versus CEC. In general, this approach is interesting and relevant as clearly one single PNEC will not be able to cover all soils in Europe. A normalisation of the PNEC implies that the normalisation equation in fact relates to a specific organism. A crucial point is then for which species the models are derived and whether the models can be applied to other species. The underlying assumption for applying the model to other species is that the rela-

tive toxicity between species is fixed across soil type. This should be discussed and verified.

MERAG proposes to apply a *lab-to-field factor* to the toxicity data. The factor is based on a comparison between the toxicity for an organism when exposed in spiked soil and in aged/field contaminated soil. The comparison between two laboratories tests, the test with the spiked soil and the test with the aged/field contaminated soil is not a laboratory-to-field extrapolation as such; it is only a part of the extrapolation. In order to make a lab-to-field extrapolation, additional issues should be considered (see Annex 6). For instance, in the field, organisms are usually exposed over generations and they are exposed to additional stressor such as climate, competitive stress and mixtures. Therefore, as long as only the effect of ageing on the bioavailability of the metals in soil is covered by the lab-to field factor, another term for this factor should be found in order to avoid confusion.

Finally, from a Nordic perspective, the implementation of the 3 issues above calls for careful analysis of the results, as the underlying models and factors are based on central European conditions.

Fact sheet No 7 on uncertainty analysis

The aim of the fact sheet is to give an introduction to techniques and methods of uncertainty and sensitivity analysis, which can be used to cast light on and quantify the uncertainties in results from e.g. a risk assessment.

Generally, the Fact sheet 7 introduces quite well most of the cornerstones of uncertainty analysis without getting into too many details. However, revisions are needed to improve clarity, and more rigorous definition of some key terms, and consistent use of them in the text would be recommended. There are also several values presented in the Fact sheet, for practical application, that are debatable, dubious, or incorrect (see Annex 7). What is also lacking is a more thorough description of the different ways to categorize uncertainty and more emphasis on the fact that some types of uncertainty are very difficult to estimate quantitatively, and/or are irreducible, and that therefore all uncertainty analyses will basically remain more or less uncertain. The last point is very important to emphasize more to avoid giving the reader the impression that by uncertainty analysis one has covered and “overcome” all uncertainties in the results.

It is recommended that this Fact sheet only focuses on the way to perform uncertainty- and sensitivity analysis on a general level, and does not go into practical details on what percentiles, application factors or other particular parameter values should be selected. Instead, one should emphasize more the practical pitfalls in uncertainty- and sensitivity analysis (see Annex 7).

Fact sheet No 8 on environmental classification

The aim with the fact sheet is to develop a strategy on classification for effects on the aquatic environment of metals, metal compounds and alloys. The approach taken is similar to and partly based on agreed strategies for environmental classification of metals and metal compounds (OECD 2001 and UN 2005), e.g. a comparison of transformation/dissolution and toxicity reference values

Since globally agreed guidance for classification purposes on aquatic hazard of metals and metal compounds already exist, there is no need to include this into MERAG. Hence, fact sheet No 8 should be replaced with references to agreed guidance (OECD 2001 and UN 2005). Deficiencies in agreed guidance, if any, could be identified, but suggestions and proposals for their improvements should be brought forward to relevant bodies (i.e. OECD/UN and RIP 3.6) for further development and agreement.

ANNEX 1:

Fact sheet 1 - “Risk characterization – general aspects“.

Reviewed by: *Torsten Källqvist*

Norwegian Institute for Water Research (NIVA), Norway

Aim and structure of the fact sheet

Fact sheet 1, integrates the principles outlined in detail in other fact sheets into a risk assessment strategy. This review will focus mainly on these general principles and not the underlying concepts.

Fact sheet 1 presents “the general building stones of a risk assessment strategy that will allow compliance in an anticipative way with forthcoming legislative obligations while ensuring, at the same time, that the best option for managing the potential risks presented by metals/metal compounds is considered”. To this end a tiered approach is proposed where the risk assessment can be refined if necessary along two different lines. The first line starts with assessment based on total concentrations, where the refinements include considerations of dissolved fractions, bioavailable fractions and, finally, the metalloregion concept. The second line of refinement goes from deterministic to probabilistic risk assessment, with field validations as a further refinement.

Comments

The general concept of a tiered approach is supported as long as the first tier is based on conservative estimates of PEC and PNEC. This implies that further refinement of the risk assessment will not be required when a risk quotient less than one has been obtained.

The concept of using bioavailable fractions in estimates of PEC and PNEC will improve the confidence of risk assessments. However, the limitations of the models used (e.g. BLM) as discussed in the review of Fact sheet 5 must be recognised.

For further refinement along this line, a metalloregion approach is proposed in Fact sheet 1. The purpose is that geographical/geological as well as biological characteristics shall be considered in risk assessments on a regional or local scale. The rationale for this approach is that the biota may be adapted to regional or local variations in natural background concentrations. As discussed in the review of Fact sheet 3, the acclimatization/adaptation concept suffers from large information gaps and the implementation of this concept in the risk assessment is not straightforward. The ultimate approach to adopt the metalloregion concept would be

to produce a complete set of data on chronic toxicity to “endemic” species tested in a medium representing the local abiotic conditions as suggested in Fact sheet 1. It should be recognised, however, that this would be a major task since it would require testing of a sufficient number of “endemic” and locally adapted species to perform a SSD analysis. If the number of such species tested is lower than required to perform an SSD, a deterministic approach to calculate PNEC has to be used, and the result would probably not represent a refinement of the PNEC estimate.

The suggestion to refine the risk assessment by using a probabilistic approach is in general supported. It should be recognised, however, that an analysis based on the exposure concentration distribution (ECD) and species sensitivity distribution (SSD) may not give a complete picture of the ecological consequences as long as only the number of species affected, and not which species that are affected is considered. Effects on “keystone” species or species of particular importance from an ecological point of view must not be ignored even if 95 % of the species are protected. Such ecologically especially relevant species may be key species governing C, N or mineral cycles. Other species which may be regarded as particularly important include popular game fish (e.g. rainbow trout, salmon sp.) or certain crustacean species. Furthermore, we lack generally information to assess the effects of fluctuating physical stressors on ecosystems (e.g. temperature & humidity) and their influence on species sensitivity. The same applies to the impact of multi-chemical exposure and fluctuating exposure concentrations, of the individual chemicals, on the populations in the environment. Peak concentrations may be fatal to organisms during specific parts of the life cycle, which means that it may be necessary to analyse the actual temporal variation of exposure concentrations together with seasonal variations in sensitivity of the biota, rather than simply a statistical comparison of ECD and SSD. Variations in abiotic factors should also be included in the analysis in a way that accounts for possible correlations or relations between the different factors and the metal concentrations.

Finally it must be noted that there is no general acceptance of the direct use of the 5%-ile from an SSD as PNEC. The TGD suggests the use of an assessment factor of (1-5) for derivation of a PNEC from an SSD. This concept must also be included in the probabilistic risk assessment. Since certain effects of ecological significance, such as reduced competitive ability or changes in behaviour will not be revealed by single species laboratory testing, there will always be a remaining uncertainty in derivation of PNEC from an SSD. Information of effects from mesocosm studies will not be sufficient to completely remove this remaining uncertainty since there are differences between ecosystems and extrapolations involve uncertainty. For this reason an assessment factor of 1 can not be justified and a lower limit of 2 as the assessment factor when PNEC is derived from SSD is recommended.

ANNEX 2:

Fact sheet 2- “Exposure assessment”

Reviewed by: *Hans Borg*
Dept of Applied Environmental Science (ITM), Sweden

Comments

General

The MERAG documents are imbalanced in terms of criteria for data selection between the parts discussing exposure, and toxicity tests, respectively. Much more effort seems to have been put on the discussion of criteria for the selection of acceptable toxicity test in the effects assessment part, compared to the criteria for acceptance of data for exposure analysis and for evaluation of natural background levels of metals, e.g. QA/QC (quality assurance/quality criteria) procedures at sampling, sample handling and analyses.

Chapter 1: Introduction

Table 1. Advantages and disadvantages of exposure assessment using modelled or measured data.

Modelled data: The use of modelled data is claimed to provide possibilities to estimate the anthropogenic contribution, but a more detailed information on how that should be accomplished is not found in Ch 2.

Measured data: A possible disadvantage not mentioned is that the evaluation is hampered if the metal data have been produced with inadequate QA/QC procedures, especially when comparing different data sets.

Chapter 2: Exposure assessments using modelled data

Diffuse source emission inventory (2.2)

“Ideally, biogeochemical regions that take the ecological dimensions into account should be used instead of regions based solely on social, demographic, economical and geographical factors (e.g., countries, states). Different background concentrations and bio-availability corrections can then be used in correspondence with different biogeochemical regions. In practise, this may not always be feasible”.

Yes, the use of biogeochemical regions would be the ideal situation but because of lack of data it is generally not possible. If such an approach is possible in certain areas, the use of catchments based regions

for the aquatic environment, as proposed in the EU-WFD, would be preferred.

2.3.2 Data gathering and evaluation

- *Annex 1, Questionnaire for the collection of site specific data.*

It is advised to use the 90th percentile for both effluent and receiving waters. However, the 90th percentile may exclude elevated (maximum) values, which could be of ecotoxicological significance in a recipient. It would be more appropriate to use the 95th percentile, and that would also be more consequent comparing the statistical treatment of data suggested for the PNEC derivation in the effects assessments. Further, monitoring programs generally have too low sampling frequency to cover all peak levels of metals.

- *Monitoring of receiving waters*

The questionnaire recommends data report in mg/l for both effluent, recipient and background sites. It may be more practical to use µg/l, in view of the expected concentration levels.

Phys-chemical water quality variables are listed but there are no recommendations for their sampling frequency.

- *Sediment monitoring*

In the questionnaire, representative metal levels (mg/kg DW) upstream and downstream the plant discharge points should be reported. But, no instructions are included for sampling methods, which sediment layers should be analysed, which analytical methods should be used, etc. Background, preindustrial, deeper levels should also be included in the evaluation, to define pollution load and enrichment factors.

In the questionnaire table SEM for some elements, as well as AVS, in µmol/l, are mentioned. However, no instructions are included for methods of sampling, sample handling and analysis. The use of standardised methods here is a crucial point, especially for such operationally defined variables as AVS and SEM.

SEM and AVS should preferably not be a mandatory first step in sediment monitoring, but rather used as a second step in relevant areas, i.e. with low oxygen and high sulphur levels. In the first step, it would be more important to advise the determination of necessary sediment characteristics such as dry weight, organic content, iron and manganese concentrations.

Annex 2: Dealing with the natural background

- *Water compartment*

Page 38: “An overview of total and dissolved background concentrations in freshwater surface waters presented by Zuurdeeg et al. (1992) and is given in *Table 2*. In the absence of local- or (eco-) region-specific background levels, the values proposed by Zuurdeeg (1992) can be used as default background values in the local or regional risk characterisation of metals.”

The data presented in *Table 2* is in general extremely high to be considered as natural background levels. The reference is an obscure unpublished report, which is not easily accessible. For example, one can compare the levels in *Table 2* with the FOREGS database for stream water (published as maps on the FOREGS website, www.gsf.fi/foregs/geo-chem/) show median values for e.g. Cd and Pb of 0,01 and 0,09 µg/l respectively, compared to *Table 2* showing mean values of 0,12 and 3,1 µg/l for these elements. The levels in the FOREGS database are then much lower, in spite of the fact that the data cannot be considered as strictly natural background levels. The sampling sites for the FOREGS database are a result of a random procedure and they rather reflect areas more or less influenced by diffuse pollution. They consequently reflect the present ambient concentrations.

The concentrations in FOREGS are in the same range as e.g. found in nationwide lake surveys in Norway, Sweden and Finland. The lakes from the northernmost parts of the countries are only moderately influenced by long-range airborne pollutants. The metal concentrations in these areas may thus be considered to be in the natural range or at least very close to that and the levels of e.g. Zn, Cu, Cd and Pb are in the same range or slightly lower than what was found in the stream waters included in the FOREGS data (Skjelkvåle et al., 2001). The data from the lake surveys may be used to evaluate the background levels of some elements in European freshwaters.

Consequently, the values presented in *Table 2* cannot be considered to represent natural background levels and should not be used for that purpose.

The “natural” background is not accurately defined in the fact sheets. It is used alternately with ambient concentrations and background concentrations at different places in the documents. Before any definite decision on how to deal with natural background concentrations, a thorough evaluation is needed of how data for that purpose should be selected and a more strict definition of “natural” should be presented.

2.4 Derivation of the Predicted Environmental Concentration using exposure models

Here is referred to a refined fate and transport model for metals, “Unit Word Model”. It is stated that the model has the capability of evaluating speciation as well as bioavailability and toxicity and that the model is currently under development. Consequently, the capabilities of the model

still remains to be proved and it seems too early to include models not even developed, validated and published in a guideline like the MERAG.

Chapter 3: Exposure assessment using measured data

3.2 Data selection and handling

The used databases for defining PECs and background levels in the MERAG and VRAR documents are not quite representative for the European range of water quality as a whole. Referring to e.g. pH and hardness, thousands of lakes show values below the 10 percentiles and will therefore be neglected in the risk assessments (cf. Fig 2 and 3, distribution of pH and Ca in European stream waters). There are several databases available today including monitoring data from numerous lakes and streams e.g. in the Nordic countries, which could be added to the risk assessment procedures in order to obtain a representative water quality covering the whole Europe (cf e.g.: www.ma.slu.se, for Swedish data).

For metals in water, it is recommended to use the dissolved fraction. That is in this case defined as the fraction passing a 0,45 µm filter. It should be pointed out however, that filtration through 0,45 µm allows passage of fine particulate, colloidal matter, which may be very important as metal carrier. Thus, the filtrate does not represent the truly dissolved fraction of metals.

Further, filtration procedures are always a potential source of contamination, sometimes causing higher metal concentrations in the filtrate than in the total water sample.

Appropriate quality control procedures and ultra clean techniques are essential to produce accurate data. Such recommendations are missing in this guideline, but should be included, or at least a reference to some other document with a detailed QA/QC protocol.

In Fig 6 and outline of the handling of data below the detection limit is presented. It is however questionable to allow as much as 90% of the data for site-specific concentrations to be below the detection limit and still try to use the data set to obtain a 90 percentile value.

Sediment and soil

In the general guidance of how to use and select monitoring data it is commented on the different fractions of metals obtained when using different digestion methods. For sediment and soil samples, the aqua regia digestion is recommended. However, there are disadvantages with aqua regia which should be mentioned. The content of hydrochloric acid may cause some interference problems when using ICP-MS and GFAAS for the metal determinations.

Apart from the soil standard method mentioned in the text (ISO 11466), using *aqua regia*, there are recently implemented international

standard methods using nitric acid as well (EN-ISO 15586 and EN-ISO 15587).

Numerous data on metals in sediments have been produced since the 70s by using nitric acid digestion, e.g. in the Nordic countries, because of the existence of valid national standard methods from the 70s and 80s (Rognerud & Fjeld 2001, Johansson 1989, www.ma.slu.se).

3.3.2 Diffuse ambient concentration exposure assessment

The seasonal variation of metals in especially running waters could be considerable. The influence of seasonal variations on sampling programs is only briefly discussed. More detailed instructions for annual sampling frequency and sampling seasons should be included in the guidelines for the determinations of ECD, PEC and in the criteria for data selection.

It is stated that diffuse ambient PEC_{area} should not be affected by the anthropogenic input of nearby point sources. However, if the measured ambient PEC is used as PEC_{regional} in calculations of site specific PEC_{local}, the contribution from the other point-sources in the region should be included in the PEC_{regional}.

Historical contamination

The principal of not using data influenced by historical contamination has been adopted in 793/93/EEC. However, it is of course important to evaluate the possible influence of historical contamination when performing risk assessments in regions influenced by such activities. The problem is often to recognise that and separate it from other diffuse sources or natural variation of metal concentrations.

It is stated in Annex 2 that the historical pollution in many cases can not be distinguished from the natural background concentration. That is generally true, but a special case where there are good possibilities to define the anthropogenic influence is lead. The fact that the stable isotopes (²⁰⁴Pb, ²⁰⁶Pb, ²⁰⁷Pb, ²⁰⁸Pb) show different ratios in lead of different origin has successfully been used to separate the influence of historical and recent pollution of lead from the natural component in e.g. lake sediment profiles, peat, soil and teeth and bone tissues. The mostly used ratio in those cases is ²⁰⁶Pb/²⁰⁷Pb (e.g. Renberg et al., 1994, Brännvall et al., 2001). As lead is a common feature of complex sulphide ores, the isotopes may be used as a marker of historical mining activities, focused also on the processing of e.g. silver, copper and zinc.

Conclusions

- Acceptance criteria for the data used for exposure analysis and background levels of metals should be discussed.

- Quality assurance procedures need to be addressed; sampling, sample handling and analyses, and the need for ultrapure techniques.
- Seasonal variations of metals in water and how it influences the sampling strategy should be discussed and a detailed instruction should be included in the guideline.
- To be consequent, the 95 percentile in stead of the 90 percentile, should be used for metal concentrations in water from monitoring data, in line with the selection for PNEC derivation.
- Determination of AVS is recommended in the questionnaire in Annex 1. If that should be included in the questionnaire, it must be accompanied by detailed instructions for sampling etc., in order to obtain any reliable data.
- The “natural” background levels of metals in water in Table 2 of Annex 2 cannot be used to define a natural background level, as they are extremely high compared to other data.
- A clear definition of natural background and the criteria for data selection should be included.
- For the digestion of sediment and soil samples, only aqua regia is recommended. The disadvantages of that need also to be commented. Standard methods using nitric acid are also available, and numerous metal data have been reported using nitric acid as well.
- The selected data for MERAG and VRA documents are not covering the total range of water quality in Europe, as many water systems are below the defined 10 percentile, e.g. regarding hardness and pH. Monitoring data covering such areas are available and should be included to obtain a less biased dataset.
- Historical pollution may be evaluated in some cases using ratios of stable lead isotopes. In some areas, lead isotopes may act as a general tracer of sulphide ore mining in the past.

References

- Brännvall, M.L., Kurkkio, H., Bindler, R., Emteryd, O., Renberg, I., 2001. The role of pollution versus natural geological sources for lead enrichment in recent lake sediments and surface soils. *Environ. Geol.* 40, 1057-1065.
- EN-ISO, 2003. Water quality - Determination of trace elements using atomic absorption spectrometry with graphite furnace. International Organization for Standardization, EN-ISO 15586.
- EN-ISO, 2002. Water quality - Digestion for the determination of selected elements in water - Part 1: Aqua regia digestion. International Organization for Standardization, ISO-15587-1.
- EN-ISO, 2002. Water quality - Digestion for the determination of selected elements in water - Part 2: Nitric acid digestion. International Organization for Standardization, ISO-15587-2.
- ISO, 1995. Soil quality - Extraction of trace elements soluble in aqua regia. International Organization for Standardization, ISO-11466.
- Johansson, K., 1989. Metals in sediment of lakes in northern Sweden. *Water, Air, and Soil Pollut.* 47, 441-455.
- Renberg, I., Wik-Persson, M., Emteryd, O., 1994. Pre-industrial atmospheric lead contamination detected in Swedish lake sediments. *Nature* 368, 323-326
- Rognerud, S., Fjeld, E. 2001. Trace element contamination of Norwegian lake sediments. *Ambio* 30, 11-19.
- Skjelkvåle, B.L., Andersen, T., Fjeld, E., Mannio, J., Wilander, A., Johansson, K., Jensen, P., Moiseenko, T., 2001. Heavy metal surveys in Nordic lakes: concentrations, geographic patterns and relation to critical limits. *Ambio*, 30, 2-10

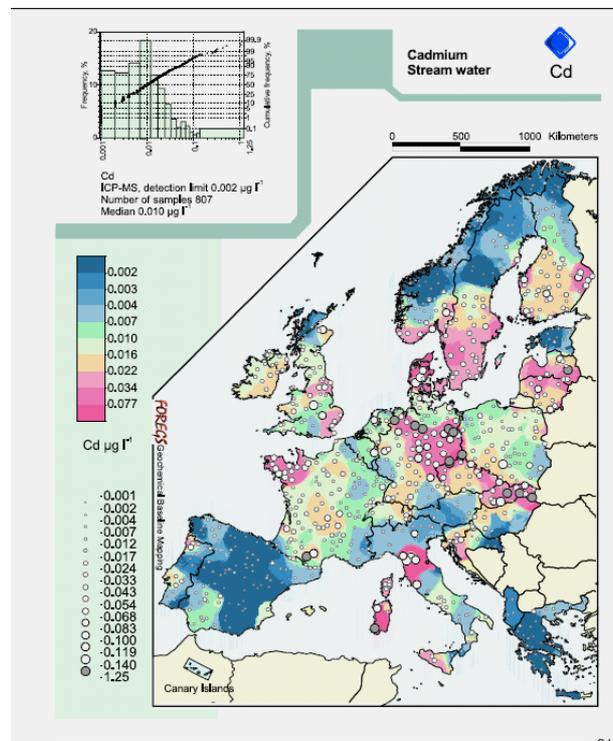


Fig.1. Cd in European stream waters, median value 0,010 µg/l (www.gsf.fi/foregs/geochem).

Fig. 2. Distribution of pH in European stream waters (www.gsf.fi/foregs/geochem).

Fig 3. Distribution of Ca in European stream waters (www.gsf.fi/foregs/geochem).

ANNEX 3:

Fact sheet 3 “Data compilation, selection and derivation of PNEC values for the risk assessment of different environmental compartments (water, sewage treatment plants, soil, sediment)” and Chapter 1 in the background document: “Metals are intrinsic parts of the environment”

Reviewed by: *Göran Lithner*,
Dept of Applied Environmental Science (ITM), Sweden

Aim with the fact sheet

”Data selected for PNEC derivation should comply with the requirements (criteria) for data quality and data relevance taking into account metal specific considerations. Therefore, it is deemed necessary to develop first a set of metal specific reliability and relevance criteria against which to evaluate the ecotoxicity data to be used. In the subsequent sections guidance is provided on data quality, aggregation, interpretation, derivation of the PNEC values”

Comments and suggestion

Data compilation and selection (chapter 2)

Physico-chemical test conditions (2.1.1.3)

It is stated in table 2 that “physico-chemical parameters which should preferably be reported, are; temperature, oxygen, hardness, salinity and pH”.

We suggest that levels of macronutrients are reported, especially in artificial media used for testing of algae, higher plants and micro-organisms.

Chemical analysis of test water (2.1.1.4)

It is stated that “there is a preference for using measured data. Nominal conc. could be considered as long as soluble salts have been used and the reported affect levels are well above the background. However if the effect levels are close to reported metal background concentration only measured values should be used “:

We agree that measured values are better to use than nominal values and that the need to use measured values increases with decreasing concentrations. However, the criterion for requirement of measured concentration in case the effect concentration is close to the background concentration “ $\mu_{\text{nominal}} - 1.95 * \sigma_{\text{nominal}} \leq \mu_{\text{background}} + 1.95 \sigma_{\text{background}}$ ” (p.5), is not

understood, because if the effect concentration is a nominal NOEC, a standard deviation of this value can not be calculated.

Much attention is in MERAG paid to the secure the quality of toxicity data, while nothing is said about the quality of measured chemical data, although erroneous data may lead to wrong conclusions. Thus, if concentrations of metals in control waters are too high (due to contamination of samples or inaccurate analysis) this may lead to the wrong conclusion that LOEC is lower than the background level (cf. p.6), or it may falsely indicate acclimatization of test organisms.

Chemical data should therefore be carefully scrutinised before being accepted. An urgent issue is therefore to elaborate objective selection criteria with the aim to guarantee quality and representativity of chemical data in laboratory and field.

Conclusion: Criteria selecting for quality and representativity of analytical data need to be elaborated.

Concentration-effect relationships (2.1.1.5)

Page 6 (last sentence): "If the resulting L(E)Cx value should be below the lowest applied control level (background level) their reliability/relevance should be evaluated." It is unclear how this could happen.

Consider the possibility that concentration-effect relationships for essential metals (esp. Cu and Zn) may have a curved shape with a "threshold".

Biological relevancy of end-point (2.1.2.1)

"For microbial testing in soils, the use of NOECs based on functional parameters (e.g. C- and N-mineralization) are preferred over NOECs based on enzymatic processes"

We agree on the importance of measuring functional parameters in soils. However, we consider that NOECs from effects on enzymatic processes may also be relevant for use in risk assessment. In addition, we consider that functional parameters are of importance for aquatic systems as well. For example, the decay of leaf-litter is an important source of macronutrients (C, N, P) in running waters, that is open for measurements (e.g. Bonada et al. 2006) (see next page).

Conclusion: For the soil compartment, enzymatic parameters should not be disregarded.

Relevancy of the species (2.1.2.3)

"For preliminary, screening-level risk assessment purposes in general both endemic and non-endemic species are usually considered relevant".
 "The species assemblage should focus on representative species for the different compartments: i.e. primary producers, primary consumers and secondary consumersfor the aquatic ecosystem" To these criteria it was added in the VRAR for copper that "used test media should be repre-

sentative for the environmental compartment(s) studied” (Table 3-2, p. 10). Note that measurements of decomposers are considered for soils and sewage treatment plants, while they are not considered for aquatic ecosystems.

Observe that biological tests conducted in nutrient rich media seldom mimic conditions in natural waters. It has been found that high levels of macronutrients N, P may interfere with metal toxicity in algal tests by suppressing it. Quite another aspect of macronutrient interference is that surplus of metals (e.g. Zn) may induce P-deficiency in algae (Kamaya et al. 2004, Guasch et al. 2004, Rijstenbil et al. 1998; Paulsson et al. 2000, Paulsson et al. 2002).

Conclusions:

1. Metal toxicity in algae may be underestimated if nutrient rich test media are used.
2. In running waters, effect parameters measuring the decay of leaves should be considered (cf. 2.1.2.1 above).

Relevancy of exposure duration (2.1.2.4)

It is stated in the fact sheet that “both acute and chronic data can be used for derivation of PNEC values, (but that) preference should be given to the use of chronic data if available.” An overview of standard toxicity tests used for classification purposes is presented in fact sheet 8, appendix A. In the text is mentioned *Ceriodaphnia* (7 d), *Daphnia* (21 d), fish (30 d), *Hyalella azteca* (28-42 d) and algae (3 d).

However, not all critical stages or “bottle necks” are covered by toxicity data. For example is the smoltification stage of salmon particularly sensitive to delayed effects of episodic exposure to aluminium (Rosse-land, pers. comm.).

The fact that most chronic toxicity experiments have a duration < 1 month, have implications for the interpretation of data from monitoring programs, that are further dealt with in chapter 3.1, under episodic exposure.

Acclimatization/adaptation (2.1.2.5 in the fact sheet, and 1.2 in the background document)

It is stated in the fact sheet that “ due to the ubiquitous presence of metals in the natural environment, organisms have been conditioned to these back-grounds since they have evolved in the presence of natural background concentrations. This theory is applicable to all metals and is even more crucial for essential metals (EE).“ It is proposed that “exposure of organisms to the natural background level reflects in fact the theoretical lower limit of the predicted no effect concentration (PNEC), i.e. a concentration, which from an evolutionary perspective, does not present a risk to the survival to the species”. In the background document it is pro-

posed that the 90th or 95th percentile of the range of natural background concentrations in the region of interest could be used as a reasonable lower limit for the PNEC.

The above mentioned make little sense as far as the issue “current/natural background levels” has not been critically reviewed (cf. 1.2.5). The risk for underestimating PNEC is obscured by lack of information in the fact sheet. The concept suffers from lack of substantial information. Since background levels are not objectively defined it can not be stated if organisms are subject to stress at “current” background levels, or if current level is a true base-line.

It is stated in the background document that: “In general the *window* of essentiality is situated in the typical range of bioavailable natural background concentrations” (p. 6). Arrows in *figure 1* indicate that the typical range of natural background concentrations is broader than the optimal concentration range for essential elements (OCEE), but this is not supported by any factual evidence. So the curves are hypothetical and requires for substantial information on actual background levels and deficiency levels for different species under different conditions. Such information may not be available.

To illustrate this with a concrete example, let us consider Knauer et al. (1997) who studied effects of Cu and Zn on growth of limnetic green alga in nutrient rich OECD medium. They found symptoms of copper deficiency at 10^{-15} - 10^{-16} M (0.1-0.01 pg/L) expressed as Cu^{2+} , and zinc deficiency below 10^{-12} M (0.1 ng/L) expressed as Zn^{2+} . They concluded that copper might have a limiting role for the growth of certain algal species in *eutrophic* lakes.

This example shows that the biological demand for essential metals may be full-filled at extremely low concentrations of free metal ions in the medium. This again illustrates the need to consider other modifying factors as macro-nutrients and productivity² (cf. 1.2.2).

Conclusions:

1. The acclimatization/adaptation concept suffers from large information gaps. There is need for objective information showing the factual ranges of background concentrations in selected areas, and factual OCEEcurves for selected species of aquatic organisms.
2. Metal deficiency (and toxicity) in aquatic organisms need to be concretised and illustrated by importance of macronutrients and other modifying factors.
3. In the background document (Chapter 1) there is a strong emphasis on essentiality. This may obscure the fact that many metals are not

² For commercial reasons, in aquacultures (especially fish farming) much attention has been paid to the optimal range of essential metals (e.g. Zn, Cu), as a means to optimise the productivity of these systems; while the optimal range in oligotrophic and ultratrophic systems is not very well studied

essential. Therefore some supplementary information on this issue should be added.

The shortage of information may hamper the implementation of principles/ methods suggested in the fact sheet, e.g. that:

“The background concentrations in culture media ideally should be representative for the organism and area under observation”, or “The essential metal concentration in culture medium should be at least equal to the minimal concentration causing deficiency for the test species”.

The acclimatization/adaptation concept is closely related to “biogeochemical region” or “metalloregion” concept mentioned below.

”Metalloregion” (2.1.2.5 in the fact sheet and 1.3 in the background document).

According to the background document, “the metalloregion concept arises from the fact that different ecoregions can be identified based upon soil type, climatic factors, latitude and elevation. Within ecoregions, sub-ecoregions (also called “metalloregions”) can be differentiated based upon natural background concentration of the metal under consideration and the presence of well-defined abiotic factors that influence metal bioavailability”. “Apart from metal background concentrations, different controlling factors must receive consideration (e.g. Ca, Mg, pH, buffering capacity for water...)”. It is further stated that “metalloregions can be developed relatively *easily* as further refinements of existing ecoregions”.

The text in this chapter is full of reservations and is not supported by scientific evidence in the text³. From the scarce information in the text it is hard to see in what way background levels in surface waters depend on other factors. Thus, it is unclear if due consideration is taken to factors as (i) Hydrological regime, (ii) Biogeochemical processes in groundwater/soil/surface water interfaces, and (iii) Turnover of metals in surface waters (lakes and rivers).

If all relevant factors are considered, the aquatic system approach may be extremely complicated and not so easy. A classification based on water type may be more relevant and more easy to implement on basis of existing data⁴. This might help to visualize the spatial distribution of lakes and rivers in the EU, and to stepwise describe basic water chemical parameters (e.g. hardness, DOC, pH) in representative water-bodies.

Conclusions:

1. The metalloregion concept has to be explained much more in detail.

³ The interpretation of « metalloregion » and related concepts (1.2.4 and ACA) is hampered by the fact that references cited in the text could not be found in the reference list or in peer-reviewed journals. Information not accessible for review was: Zuurdeg (1992), Van Asche et al. (1997), Janssen et al. (2000), Muysen (2002), and Fairbrother & McLaughlin (2002).

⁴ Use e.g. FOREGS, EU (Corine, and CCM river and catchment database), and National data bases, and data produced in the European freshwater monitoring network.

2. For the aquatic compartment, elaborate a classification based on water type.
3. Visualise the spatial distribution of lakes and rivers and some basic characters (e.g. hardness, DOC, pH). Start with FOREGS data as a surrogate for the EU rivers.
4. Use this information to increase the realism of forthcoming “worst case” scenarios.

Derivation of the predicted no effect concentration (Chapter 3)

Aggregation/selection of L(E)C50/NOEC data (3.1)

It is stated that : “If for one species more than one L(E)C50 /chronic NOEC values based on the same toxicological endpoint are available, these values are averaged by calculating the geometric mean”.

However, a geometric mean can only be calculated if the tests are carried out under similar conditions (or it can be proven that the difference in conditions does not affect the outcome of the test). In the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) it is recommended that the geometric mean can be used when 4 or more values are available for the same species. When less than 4 data are available, the lowest acceptable value is generally used.

Approach for the derivation of PNEC-values in water (3.2)

It is stated that: (i) ”When *limited* number of data are available, PNEC setting is based on the use of assessment factors...”. (ii) If *sufficient* ecotoxicological data are available the use of a statistical extrapolation method is recommended”. It is further stated that “Once a distribution is selected, the PNEC value is derived as the 5th percentile of the SSD. The use of *an additional assessment factor on this value is not recommended*” (3.2.3.1, p. 14).

Ecological relevance (3.1, p.10; 3.2.3.2, p.14)

(i) “Our emphasis is on the representativeness of all taxa that could be relevant for the functioning of the ecosystem” (p.10, footnote 10)⁵

(ii) “Data should be representative in terms of ecological relevance (e.g. include key species) and include the appropriate number of taxonomic groups and trophic levels” (p. 14)

Specific requirements to be fulfilled before using statistical extrapolation methods (3.2.3.2) are that: ”Available data should be representative in terms of ecological relevance and include the appropriate number of taxonomic groups and trophic levels”. It is suggested that: ”The Species Sensitivity Distribution (SSD) in water should cover at least 8 taxonomic

⁵ All taxa are per definition relevant for the functioning of the ecosystem.

groups containing at least 10 NOECs (preferably >15) for different species (Table 3)”.

It is questionable if the suggested procedure (at least 8 taxonomic groups and 10 NOECs; preferably > 15) will fulfil the ecological relevance criteria mentioned above, as far as the issue, *ecological relevance*, is not defined and considered for its implications. Furthermore, this procedure only considers variations (in sensitivity) among taxonomic groups, while not considering variations among taxa from the same taxonomic group (e.g. *Baetis* vs. *Aeschna*, *Gammarus* vs. *Asellus*).

Neither is the meaning of “key species” defined in the text, but might be about equivalent with keystone species. Whatever term is used it would mean that the species in question is critical for the structure and functioning of the ecosystem. For example, *Gammarus pulex* and some bivalves are well founded keystone species due to their structuring effect on aquatic ecosystems.

Sensitive key species or keystone species may therefore be more important than other equally sensitive species. They should therefore be given extra weight in the risk assessment.

Conclusions:

1. We suggest that any sensitive keystone species and ecosystem processes should be considered, by increasing their weight in derivation of PNEC and/or assessment factor.
2. If necessary the database and the ecological framework of MERAG should be improved to full-fill these goals.

The consequences of not full-filling relevance criteria for data rich substances are discussed more in detail in the next chapter.

Assessment factor and related issues in water

According to TGD an extra assessment factor of between 5 and 1 should be applied to derive the PNEC from HC5. This should reflect the further uncertainties regarding for instance if all sensitive life stages, all taxonomic groups and life forms and differences in sensitivity between lab and field are covered. Such uncertainty is not covered by an SSD or its confidence intervals.

The MERAG statement, that no assessment factor is recommended for data-rich substances, is resting on several assumptions, with negative ecological implications if not full-filled.

First, alternative exposure routes are to great extent ignored by MERAG, although adverse effects of diet-borne metals have been reported in fish and planktonic crustaceans exposed to metals in natural food items (e.g. Hansen et al. 2004, mixed pollution; De Schampelaere et al. 2004, Zn; Hook & Fisher 2001, Ag; Sofyan et al. 2006, Cu; Bielmeyer et al. 2006, Cu, Zn, Ni or Ag). In addition, it is well known that copper

salts added to pelleted food may hamper the growth of juvenile Rainbow trout/Atlantic salmon at 460-1000 µg Cu/g DS (reviewed by Clearwater et al. 2002)⁶.

Second, any exposure shorter than 1-2 months at concentrations exceeding the 90th percentile of monitoring data on environmental concentrations, is not considered in the risk assessment procedure suggested by MERAG. This is problematic, since many NOEC-values are derived from experiments with duration of one month or less, and from our experience, the sampling interval in most monitoring programs exceeds one month. Therefore, most monitoring programs will fail to register any additional risks of short-term exposure to concentrations higher than the 90th percentile.

This problem may be illustrated by data from Norway showing delayed mortality in caged salmon fry after short-term exposure (< 1 mo) to aluminium. This mortality continued for 1-2 weeks after the concentration of Al in fish gills had returned to normal level (Teien et al. 2005). During this episode, the concentration of inorganic Al in water increased by a factor 2-3 compared with the base-level. This example also illustrates that risks for acute effects of episodic exposure may be especially high for metals with low acute/chronic ratios, such as copper (and zinc?), often having acute/chronic ratios as low as 3:1.

Third, every metal is considered isolated and not in combination with other metals.

Fourth, BLM is used by MERAG as a generic tool for incorporation of bioavailability and normalization of toxicity data. However, existing models are still not fully validated. In addition, deviations from BLM already have been documented e.g. in adult bivalves, accumulating metals not only from water but also from algae, bacteria, colloidal matter, and truly dissolved (< 1 kD) organic matter (Guo et al. 2001, Lorenzo et al. 2005, Roditi et al. 2000, Croteau et al. 2005). But it remains to show if exposure from additional up-take is affecting the fitness of bivalves in the same way as fish and crustaceans (cf. p. 6).

Fifth, the need for an assessment factor increases with the geographic extension of the polluted area ("scale factor"). Therefore, impact of large-scale pollution affecting sea and great lakes, deserves higher assessment factors than does impact of small scale pollution.

Conclusions:

- For data rich substances, assessment factors (>2) should be applied to PNEC, in order to compensate for (i) additional exposure routes, (ii)

⁶ It may be argued that these concentrations are higher than found at 9 µg Cu/l (being the PNEC of the suggested "worst case" scenario in VRA Cu). However, even higher concentrations may occur in certain food items, e.g. isopods. Furthermore, concentrations are not directly comparable, because efficiency of Cu uptake is much higher from natural food items than from food pellets (see e.g. Clearwater et al. 2002, Croteau et al 2005).

episodic exposure, (iii) combination effects, (iv) uncertainties of used speciation models, and (v) the “scale factor”.

- Sensitive keystone species and ecosystem processes should in first hand be considered by increasing their weight in the derivation of PNEC-values (final step/weight-of-evidence?).
- For proper handling “worst case” scenarios and episodic exposure (< 1 mo), we suggest that 10th /90th percentiles are replaced by 5th/95th percentiles to increase realism and use similar statistical criteria for protection of sensitive water-bodies as for protection of sensitive species.
- To compensate for large collective risk an extra assessment factor, the “scale factor”, could be applied to large water-bodies, e.g. sea gulfs, great lakes and open sea. We suggest a graded approach where AF follows the target area along a logarithmic scale.

Approach for the derivation of PNEC-values in soil and sediment (3.2.3.2, Tabel 4, p. 16)

”For the sediment compartment in particular, a more pragmatic approach should be developed based upon expert judgement since the number of taxa for which internationally accepted test protocols are available is limited”. Recommendations in table 4 are that ”the SSD should cover three taxonomic groups (minimum) and contain at least 4 NOEC values for different species, It is also recommended that (three) different living and feeding strategies should be represented”.

“For soils both the tests on terrestrial species (plants and invertebrates) and the tests on microbial function can be used to derive the SSD for the terrestrial compartment “. In table 4 is recommended that “The SSD should cover at least three taxonomic groups, one of which is a plant and contain at least 4 NOEC values for different species”. These recommendations, in relation to the use of SSDs for the sediment and soil compartments, were discussed at the Oslo workshop. The recommended numbers of taxonomic groups were considered to few, to sufficiently cover the range of organisms in respective compartment. The workshop came up with the following *conclusions*:

- For the sediment compartment, a minimum of six taxonomic groups should be required otherwise the assessment factor should be increased.
- For the soil compartment, secondary consumers (e.g. mites) should be added as an additional taxonomic group (see also 2.1.2.1)

Effects weight-of-evidence (3.2.4)

“Once a PNEC has been derived for the different compartments, either through use of assessment factors or by using a statistical extrapolation

method, the weight-of-evidence from all other available data including mesocosm/field data etc... should be evaluated in a final tier”.

“ If the results of laboratory and field (model) ecosystem studies show that the effects on ecosystems are unlikely to occur at the derived PNEC level, the assessment factor may be decreased. On the other hand, if it is clear that the PNEC is under-protective for the ecosystem the mesocosm value could be used”.

It is suggested in the fact-sheet that the most important criteria to be full-filled are: (1) distinct concentration-effect relationship, (2) a reliable multi-species NOEC, (3) duration >28 d, (6) measurement of test compound concentrations, (7) measurements of bioavailability modifying factors.

Conclusions:

- We agree that results from mesocosm studies are to be considered in the final risk assessment procedure, if they cover a relevant concentration interval.
- In addition to total biomass, species diversity, and species richness, any effects on individual species could be included. Special attention should be paid to *functional parameters* e.g. measurements of decay of litter, gross respiration, gross primary production and total filtration capacity (filter feeders).

References

- Bielmeyer, G. K., Grosell, M. & K. V. Brix (2006), Toxicity of silver, zinc, copper, and nickel to the copepod *Acartia tonsa* exposed via a phytoplankton diet. *Environ. Sci. Technol.* Vol.40: 2063-2068.
- Bonada, N., Prat, N., Resh, V. H. & B. Statzner (2006), Development of aquatic insect biomonitoring: a comparative analysis of recent approaches. *Annu. Rev. Entomol.* Vol 51: 495-523.
- Clearwater, S.J., Faraq, A. M. & J. S. Meyer. (2002), Review. Bioavailability and toxic-ity of dietborne copper and zinc to fish. *Comp. Biochem. Physiol. Part C* 132, 269-313.
- Croteau, M-N. & S. N. Luoma (2005), Delineating copper accumulation pathways for the freshwater bivalve *Corbicula* using stable metal isotopes. *Environ. Toxicol. Chem.* Vol 24 (11): 2871-2878.
- EC (2003). Technical Guidance document on Risk Assessment *in support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market.* European Commission, Joint Research Center, Institute for Health and Consumer Protection, European Chemicals Bureau, Ispra (VA), Italy.
- De Schampaelere, K. A. C, Canli, M., Van Lierde, V., Forrez, I., Vanhaecke, F. & C. R. Janssen (2004), Reproductive toxicity of dietary zinc to *Daphnia magna*. *Aquat. Toxicol.* Vol 70: 233-244.
- GHS (2003), Globally Harmonized System of Classification and Labeling of Chemicals, UNECE, ISBN:92-1-216463-3
- Guo, L., Hunt, B. J., Santchi, P. H. & S. M. Ray (2005), Effects of dissolved organic matter on the uptake of trace

- metals by american oysters. *Environ. Sci. Technol.* Vol 35: 885-993.
- Guasch, H., Navarro, E., Serra, A. & S. Sabater (2004), Phosphate limitation influences the sensitivity to copper in periphytic algae. *Freshw. Biol.* Vol 49: 463-473.
- Hansen, J. A., Lipton, J., Welsh, P. G., Cacula, D. & B. MacConnell (2004), Reduced growth of rainbow trout (*Ochorhynchus mykiss*) fed a live invertebrate diet pre-exposed to metal contaminated sediments. *Environ. Toxicol. Chem.* 8, 1902-1911.
- Hook, S. E. & N. S. Fisher (2001), Sublethal effects of silver in zooplankton: importance of exposure pathways and implications for toxicity testing. *Environ. Toxicol. Chem.* Vol 20 (3): 568-574.
- Kamaya, Y., Takada, T. & K. Suzuki (2004), Effects of medium phosphate on the sensitivity of *Selenastrum capricornutum* to chemicals. *Bull. Environ. Contam. Toxicol.* Vol 73: 995-1000.
- Knauer, K., Behra, R. & L. Sigg (1997), Effects of free Cu^{2+} and Zn^{2+} ions on growth and metal accumulation in freshwater algae. *Environ. Toxicol. Chem.* Vol 16(2): 220-229.
- Lorenzo, J. I., Beiras, R., Mubiana, V. K. & R. Blust (2005), Copper uptake by *Mytilus edulis* in the presence of humic acids. *Environ. Toxicol. Chem.* Vol 24(4): 973-980.
- Paulsson, M., Nyström, B. & H. Blanck (2000), Long-term toxicity of zinc to bacteria and algae in periphyton communities from the river Göta Älv, based on a microcosm study. *Aquat. Toxicol.* Vol 47: 243-257.
- Paulsson, M., Månsson, V. & H. Blanck (2002), Effects of zinc on the phosphorus availability from the river Göta Älv. *Aquat. Toxicol.* Vol 56: 103-113.
- Rijstenbil, J. W., Dehairs, F., Ehrlich, R. & J. A. Wijnholds (1998), Effects of nitrogen status on copper accumulation and pools of metal-binding peptides in the planktonic diatom *Thalassiosira Pseudonana*. *Aquat. Toxicol.* Vol 42: 187-209.
- Roditi, H. A., Fisher, N. S. & S. A. Sanudo-Wilhelmy (2000), Uptake of dissolved organic carbon and trace elements by zebra mussels. *Nature* Vol: 407-78-80.
- Rosseland, B.O., Blakar, I.A., Bulger, A., Kroglund, F and Kvellstad, A. (1992). The mixing zone between limed and acidic river waters: Complex aluminium chemistry and extreme toxicity for salmonids. *Environ. Pollut.* 78 (1-3), 3-8.
- Teien et al. (2006), Gill reactivity of Al-species following liming. *Sci. Total Env.* 358(1-3), 206-220.
- Sofyan, A., Shaw, J. R. & W. J. Birge (2006), Metal trophic transfer from algae to clado-cerans and the relative importance of dietary metal exposure. *Environ. Toxicol. Chem.* Vol 25(4): 1034-1041
- Magna, Sed et Victoria Nee (2004), Quisque non nulla. *Integer eu dolor id* vol 33 (4): 401-404

ANNEX 4:

Fact Sheet 4 - “Marine risk assessment” and chapter 4 of Background Document “Biodiversity and sensitivity of marine versus freshwater systems”

Reviewed by: *Ketil Hylland*

Norwegian Institute for Water Research (NIVA), Norway

1. Introduction, aims and structure of the fact sheet

Marine risk assessment was included in the TGD in 2003 (EC, 2003). The TGD comprises guidance on exposure assessment, effects assessment and PBT assessment on elements specific to marine environments.

Three documents were available for this review:

3. Fact sheet 4: Marine Risk Assessment
4. “One-pager” 13: Use of Freshwater ecotoxicity data to determine a marine PNEC
5. Metal risk assessment guidance document: Background Document (section 4).

There is a large extent of overlap between the three documents. Although the specific aims for each document were not stated explicitly, they could be extracted as follows from the documents.

1.1. Aims

The aims of the fact sheet appear to have been to characterise marine ecosystems (abiotic and biological components), compare species diversity in freshwater and marine ecosystems, compare species sensitivity between the two ecosystems and to develop a proposal for PNEC derivation for marine ecosystems when there is a paucity of data.

The aims of the “one-pager” were essentially similar to the last three of the fact sheet with a larger emphasis on methodology for developing PNECs and SSDs with limited marine data.

The aims of the Background Document (section 4) basically overlapped with the first four aims in the fact sheet (all except derivation of PNEC).

1.2. Structure

As mentioned above, the fact sheet consists of a (very) brief description of the marine chemistry of some trace metals, a discussion of diversity in freshwater and marine ecosystems, as well as a discussion of species sensitivity in the two habitats. The above information is used to support an argumentation that freshwater data can be used to supplement marine data in marine risk assessments.

The “one-pager” proposes a methodology to develop PNECs and SSDs for substances with limited or no data for marine organisms.

The structure of section 4 of the Background Document is essentially identical to most of the fact sheet (except the last section in the latter).

2. Comments to fact sheet 4: Marine Risk Assessment

The document consists of a general section on “Characteristics of the marine environment” that includes abiotic and biotic “factors” as well as a specific section on “Derivation of the predicted no effect concentration (PNEC) for marine species”.

There are several shortcomings of this fact sheet, the most important being:

- the presentation of abiotic factors (section 2.1) is of low scientific quality and does not correctly reflect the existing knowledge about metal bioavailability in marine ecosystems;
- the presentation of species diversity (section 2.2.1) is skewed, apparently from a desire to show that there is no great difference in the diversity of animal groups (plants are not included) between freshwater and marine ecosystems;
- the presentation of species sensitivity (section 2.2.2) is inaccurate and uses a dataset (which had not been quality assured) for one metal, lead, to suggest that there is no difference in sensitivity between freshwater and marine ecosystems;
- the suggestion of a procedure to derive a PNEC for marine species using freshwater data (section 3) is inconsistent and illogical; if there is not sufficient marine data for a risk assessment it will not be sufficient data to ensure no difference between freshwater and marine data either.

The above shortcomings are discussed in more detail below.

2.1 Abiotic factors

The text reflects a limited understanding of both the terminology and science of marine chemistry. One example is the term “hardness”, which is not commonly used for seawater. There is a very general discussion of

organic material in marine waters with no mention of seasonality or the different qualities of organic material (e.g. allochthonous and autochthonous). The speciation of different metal obviously varies and metals will associate with a range of anions, e.g. chloride, sulphate, carbonate, phosphate (see Turner et al., 1981; Millero, 2001), whereas this document only discusses chloride. Statements on lead complexation with chloride in seawater are at best contentious (cf Martell et al., 1988). Although briefly mentioned, the importance of the complex chemical and physical processes in estuaries and other mixing zones on metal speciation and bioavailability need a much more substantial discussion. There is a strange statement on "... natural background levels of (essential) metals in ..". There is of course a background level in natural (uncontaminated) seawater for both essential and non-essential metals.

2.2. Species diversity

As is correctly stated in the fact sheet, a substantial number of phyla are only found in marine ecosystems, whereas none are limited to freshwater systems. There have been major taxonomic revisions at all levels over the past decades and it surprising to see references from the 80'ies as backing for statements on the number of any taxonomic groupings in 2005. The actual number of species in some marine taxa is also not well established (Schander & Willassen, 2005; Snelgrove, 1998). It is of course known that insects is the largest taxonomic group on earth, but this is used in the fact sheet to give the impression that diversity is similar in the two habitats. This is obviously not true, except in a demagogical sense. In this context it is important to address the possibility that organisms from a range of phyla will have dramatically different physiologies and may therefore be differently affected by foreign substances. The characteristics of species from distant taxonomic groups (as found to a larger extent in marine systems) will differ to a much larger extent than between species from e.g. different insect orders. A number of phyla with very specific physiology and ecological traits are only found in marine ecosystems, such as ctenophores (comb jellyfish) and echinoderms.

2.3. Species sensitivity

As is mentioned in the fact sheet, freshwater and marine organisms differ in their physiology. The most important difference as concerns metal toxicity is the difference in ion regulation and hence potential ion pumps and other potential metal transporters in gills or skin.

It is obviously not scientifically viable to extrapolate from one study using data that were not quality assured for one metal to all other metals (as done in the fact sheet, for lead).

It is likely that marine food chains are longer than freshwater food chains due to a larger number of invertebrate trophic levels (invertebrate predators). Secondary poisoning will be at least as important in marine ecosystems as in freshwater systems; due to ecologically dominant filter-feeders such as mussels in combination with birds that almost exclusively feed on them (e.g. eider duck) and a larger number of fish-eating mammals (and birds).

2.3. Derivation of PNECs

If there is not sufficient marine data to perform a risk assessment, there will not be sufficient data to establish whether there is no difference between marine and freshwater data! The approach to determine whether freshwater and marine data may be used in conjunction is illogical. It is furthermore inconsistent to suggest that freshwater data may be used for marine assessments, but not vice versa. The reason for this suggestion is unclear.

3. Comments to “One-pager” 13: Use of Freshwater Ecotoxicity Data to Determine a Marine PNEC

This document addresses one of the issues already discussed above, but more specifically focusing on whether freshwater and marine data can be combined in a SSD approach.

There are two main weaknesses in the suggested approach:

- it assumes that there are no additional taxonomic groups in marine ecosystems that may be more sensitive than freshwater taxa (see above);
- if there is not sufficient marine data to develop a marine SSD there will not be sufficient data to ensure that there is no difference between freshwater and marine SSDs.

It is also not clear why freshwater data should be used for marine risk assessment, but not vice versa. One solution to the above is to use an application factor if freshwater data is used to derive a SSD.

4. Comments to “Metal Risk Assessment Guidance Document: Background Document (section 4: Biodiversity and sensitivity of marine vs. freshwater systems)”

This section focuses on another issue discussed in fact sheet 4, i.e. differences in biodiversity between freshwater and marine ecosystems and potential differences in sensitivity. The text is for a large part similar to that found in fact sheet 4 and the same comments apply (see above).

It is stated that “..on average food chains do not appear to be longer in marine ecosystems than in other ecosystems”; this is a very general and non-specific statement with limited scientific backing. The general understanding of aquatic ecology does not support this statement and if maintained should be backed up more extensively.

5. Conclusions

The description of the marine chemistry of trace metals is scientifically weak.

The discussion of species diversity and taxonomic variation in freshwater and marine ecosystems is biased and erroneous. This follows onto the discussion on species sensitivities in the two systems.

It is unlikely that toxicity testing with only freshwater species will be protective of all marine phyla. Only marine data should therefore be used for marine risk assessment. It is not really possible to provide an appropriate extrapolation factor should freshwater data be considered, but the factor should presumably not be below the factor (10) suggested in the TGD to be protective for marine species. The above argumentation pertains to the greater diversity of physiology in marine species and will have a general application for both organic and inorganic contaminants.

6. References

- EC. 2003. Technical Guidance document on Risk Assessment *in support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market.* European Commission, Joint Research Center, Institute for Health and Consumer Protection, European Chemicals Bureau, Ispra (VA), Italy.
- Martell, A.E., Motekaitis, R.J., Smith, R.M. 1988. Structure-stability relationships of metal complexes and metal speciation in environmental aqueous solutions. *Environmental Toxicology & Chemistry*, 7: 417-434.
- Millero, F., 2001. Speciation of metals in natural waters. *Geochem. Trans.* 2: 57-64.
- Schander, C., Willassen, E., 2005. What can biological barcoding do for ma-

- rine biology? *Marine Biology Research*, 1:79-83.
- Snelgrove, P.V.R. 1998. The biodiversity of macrofaunal organisms in marine sediments. *Biodiversity & Conservation*, 7: 1123-1132.
- Turner, D.R., Whitfield, M., Dickson, A.G. 1981. The equilibrium speciation of dissolved components in freshwater and seawater at 25°C and 1 atm pressure. *Geochimica et Cosmochimica Acta*, 45:855-881.

ANNEX 5:

Fact sheet 5 - “Incorporation of bioavailability for water, (soils) and sediments” and background Chapter 2 in the background document: “Bioavailability of metals/metal compounds”.

Reviewed by: *Göran Lithner*,
Dept of Applied Environmental Science (ITM), Sweden
and *Jens Skei*
Norwegian Institute for Water Research (NIVA), Norway

Aim with the factsheet

The degree to which metals are available to aquatic organisms varies depending on the site-specific geochemical conditions controlling metal speciation, and therefore also influence the degree to which organisms may be exposed to metals and experience adverse effects from such exposure”. In order to reduce the uncertainty and to increase the ecological relevance of the assessment both effects and/or exposure data should be normalized to the bioavailable fraction by incorporation of bioavailability concepts

Incorporation of bioavailability in fresh-water

Short description on the approach and proposed methods

In order to reduce the uncertainty and to increase the ecological relevance of the assessment it is proposed that both effects and/or exposure data should be normalized to the bioavailable fraction by *incorporation of bioavailability concepts*. One method is to estimate the true *free ionic metal fraction by using metal speciation programs (e.g. WHAM)* taking into account the presence of important binding ligands (e.g. DOC, chlorides etc.). For some metals toxicity based models from simple regression models to the more refined Biotic Ligand Models (BLM) are available or under development”.

Following from the analysis of ecotoxicity data in Chapter 1, a PNEC not corrected for any bioavailability should be derived. In case bioavailability models can be applied, this generic PNEC can be modified to a PNEC normalized to specific local or regional conditions, using *typical value for bioavailability modifying factors* (e.g. pH, DOC, hardness).

This “reference PNEC” is *normalized to the reasonable worst case conditions* which are based on 10th or 90th percentile of the bioavailability modifying factors, assumed to be representative for the EU. This “*reasonable worst case scenario*” is assumed to visualize worst conditions (i.e. maximized bioavailability) met in EU. The “*typical scenario*” on the other hand is based on 50th percentiles of bioavailability modifying factors and may illustrate an average situation.

For bioavailability correction different datasets of abiotic factors (and environmental concentrations) should be considered. As regards sampling strategy it is stated in the fact sheet (p. 2) that “*Datasets of abiotic factors and environmental concentrations should be representative for and should cover the specific area under investigation*”, i.e. local, regional (watershed), or be representative for larger entities (country, continent).

Comments and suggestion

General comment on the representativity in space

Since the issues “representativity” and “coverage” are not explicitly explained in the text, two examples are taken from ongoing work at the EU-level. The first one comes from VRAR Cu (Draft 15 May 2005). In this case data from 6 EU-countries were used for calculating “typical” and reasonable “worst case” scenarios, claimed to be representative for the EU. The fact that data sets were strongly biased makes them, per definition, not representative for the EU.

The second example is taken from the European geological mapping program of FOREGS (see <http://www.gsf.fi/foregs/geochem/>). This sampling program is based on the same statistical criteria as are used in the global mapping program. That is, randomised sampling and high coverage of the EU countries. This approach will certainly produce data spatially more representative for the EU than aforementioned datasets. Thus, the aim of this mapping program is in FOREGS own words “*to show the actual situation within Europe*”.

Conclusions:

1. Representative data sets in space may rest on randomised sampling and high spatial coverage, as practised by FOREGS.
2. FOREGS water chemical data are to prefer before biased data sets, when calculating “worst case” scenarios for the EU.
3. FOREGS data may be used as surrogates for current background metal levels in remote areas. Surrogates for natural background levels may be found in far north (National data bases).
4. However, the *definition* of current and natural background level is a complicated issue depending on several factors (fact sheet 3, 1.2.5)

Implementation of bioavailability: water (chapter 2)

Transformation to soluble fraction (2.2)

It is stated for the aquatic toxicity data that “aquatic toxicity tests tend to maximize toxicity. Hence it is assumed that for all toxicity data no additional conversion to a dissolved fraction has to be applied”

This is most certainly not true for all metals. For instance, in the VRAR on Pb (draft of February 2006) it was found from a literature review that dissolved Pb comprised between 38 and 83 % of total Pb in the tests where this relation was reported. This is important since it is proposed to compare the PNECs with the dissolved fraction of PEC.

Use of speciation models (2.3)

MERAG suggests the use of WHAM (Model V) described by Tipping (1994). It is also integrated in HydroQual’s biotic ligand model. There are some newer models which, at least theoretically, better consider the fact that dissolved organic material has several different functional groups which bind metals with different high strength (see reviews by Tipping, 1998, Gustafsson, 2001, Koopal et al. 2005). One might of course question the choice of WHAM (Model V) as a basis for risk assessments since it is an older model. But according to Jon Petter Gustafsson, who made the Stockholm Humic Model, it is a factor of minor importance in comparison with other uncertainties.

From a chemical and analytical point of view, the BLM concept based on chemical modelling for transformation of soluble metals to bioavailable species in water is good, pioneering and builds on new research. However, there is a tendency in the MERAG documents to underestimate the uncertainties of modelling. Some examples are given below.

Uncertainties (2.1.1)

Uncertainties of modelling are mentioned in the background document, e.g. “the enormously varied nature of DOM in aquatic systems presents major challenges in predicting metal complexation”. And further, “with regard to the speciation calculation, none of existing models have the capability to fully model the metal complexation in terms of full heterogeneity of natural waters, which include colloids, particulates and complexing cations to DOC like Al and Fe, as well as dissolved ligands.” (p. 22).

It may be added that the difficulties to measure concentrations of free Fe^{3+} and Al^{3+} ions at $\text{pH} > 6$ means that complexation constants for Fe (III) and Al in WHAM (Model V) and the other models, have not been calibrated for such conditions. Since Fe and Al are important competing ions, the simulated concentrations of bioavailable metals become uncertain between pH 6 and pH 8. For strongly complexing metals e.g. Cu and Pb the uncertainty may be as high as an order of magnitude. This may result

in underestimation of free metal ions (e.g. Cu^{2+} , Pb^{2+}), if competition from Al^{3+} and Fe^{3+} is not considered in the model (pers. comm., J P Gustafsson).

De Schampaelere & Janssen (2004b) reported that acute Cu toxicity to *Daphnia magna* in a Swedish lake water was much higher than predicted from BLM. This discrepancy was ascribed to unusually high levels of Fe and Al, in combination with low pH of 5.5 (cf. VRAR for Cu, p.26-27). However, even if the concentrations of Fe and Al may seem high from a northern continental perspective, they are common in Scandinavian waters, and probably also in Scotland, Ireland and certain other parts of Europe.

In the background document is stated that "future BLM developments will undoubtedly benefit from improvements in the detection techniques for determining low free metal ion concentrations ($\ll 1 \mu\text{g/l}$). Availability of such data will improve the modelling of metal binding to DOM, biotic ligands and other sites that play a role with regard to the manifestation of toxic effects". (p. 23). It is also mentioned that dietary exposure might contribute to remaining variability between observed and predicted toxicity. (p. 24); see p. 11 in this review for further details.

Few studies have attempted to compare model predictions of metal speciation to experimental in field results due to difficulties of making accurate measurements of chemical speciation directly on natural waters. In a recent collaborative study (Unsworth et al. 2006), including several laboratories, in situ measurements of Cd, Cu, Ni, and Pb were made in three freshwaters with four different speciation techniques. Permeation liquid membrane (PLM) and the Donnan membrane technique (DMT) were used under conditions where they estimate the concentration of *free ions*.

From the results of this work the authors concluded, "that WHAM and NICA donnan models can predict the distribution of of the *dominant* species of of trace metals in freshwater with reasonable accuracy. However, concentrations of *free ions*, when they represent a very small fraction of of total dissolved concentrations, as for Cu and Pb, *may be underestimated by several orders of magnitude*. These uncertainties are concerning, given the current importance ascribed to the free ion in models of biological uptake. There is an urgent need to perform further accurate speciation measurements in natural waters where the total dissolved concentrations of all relevant species can be accurately measured, to test further both models and methods to identify the causes of the discrepancy."

Extrapolation of BLM

The cladoceran test species *Daphnia magna* is often used as a model organism to illustrate the BLM concept, adaptation/acclimatization, and alternative exposure routes. From an ecological point of view, it is doubt-

ful to extrapolate from one or a handful of test species to other test species, and to endemic species and ecosystems to be protected.

It is said in the fact sheet that “Normalization of individual toxicity values across species and trophic levels using a specific bioavailability model is allowed if evidence can be provided that it can be translated easily to other organisms”.

However, there is no guidance on how such evidence is developed. How many and what species should the models be validated for? Under what conditions? Validation with field data?

It is also difficult to follow how the models are proposed to be applied in chapter 2.4. How are “realistic worst case (rwc), typical, generic and reference defined and when are they used? In the case when only a limited amount of toxicity values can be normalized, how do you know that the species that would result in the smallest correction for bioavailability is included?

Furthermore, experimental evidence indicates that DOC is not always protective, e.g. in *bivalves* (cf. comments on FS 2). Therefore, the FIA model is not the only driving factor for whole-body concentrations, but may be so in gills. However BLM is an effect-driven model, and few if any studies have hitherto focused on biological effects of diet-borne metals to *bivalves*!

BLM assumes stable systems at equilibrium. However, unstable situations are common in the environment, e.g. in mixing zones. Specific toxic effects of metals may occur under such situations, as has been shown e.g. for aluminium in estuarine waters (Rosseland et al. 1992, Teien et al. 2004).

Suggestions

1. Uncertainties of used models should be critically reviewed and evaluated as concerns the outcome of risk assessment procedures and assessment factors for different elements.
2. There is a need of further field validations of BLM and chemical speciation models in representative water types.
3. Discuss the representativity problem and how to achieve representative “worst case” scenarios for the EU and typical sub-regions/water-types. Discuss the meaning of a “typical” scenario based on a mean value without expressing the natural variation in statistical terms. Does it make any sense at all?

Incorporation of bioavailability of metals in sediments.

- Short description of the approach and proposed methods.

Quantitative assessment of bioavailability of metals, whether they are essential or not, is important in risk assessment of contaminated sediments. Many countries are in the initial phase of sediment remediation. The challenge is to make sure that the decisions are based on best available knowledge and technology and that the decisions are risk based. Knowledge of the bioavailability of contaminants in sediments is essential, but far from straight forward.

The document consists of an Introduction and a figure trying to demonstrate a holistic look at metals in water, soil and sediments in the context of bioavailability. Partitioning of metals between dissolved and particulate phases in the water column and between pore water and sediments in the bottom sediment compartment is important in terms of understanding bioavailability and risks.

- Comments, questions and suggestions to the fact sheet.

The use of SEM/AVS and the interpretation of AVS results should be used with care due to the following:

- Nearly all animals require oxygen, and what happens to metals in the reduced layer of sediment is then not of great importance. If SEM/AVS evaluations will be widely accepted they may complicate the regulatory/risk assessment system. In some cases they might allow high metal concentrations in sediments where bioavailability does exist (but is not measured). The measure will not be relevant for all ecosystems (e.g. oligotrophic lakes and rivers).
- In most cases there is a co-variance between metal concentration, AVS and OC, i.e. the difference in toxicity might as well be explained by metal concentration and OC. SEM/AVS might be relevant in some places but it might be misleading in others.
- There are large spatial and temporal variations of AVS in sediments between water courses as well as within one recipient, and it is complicated to collect representative samples. The samples should be representative for where the benthic fauna thrive.
- There are strong vertical gradients of AVS (particularly in marine sediments) and samples taken at depth in a sediment core may not be relevant with respect to assessing bioavailability and toxicity to organisms living close to sediment-water interphase. Most animals in sediments occupy or interface with the oxidized layers, and by their

burrowing activities in fine sediments often oxidize those burrows. Almost any value for AVS is possible to obtain by adjusting the sampling procedure at such sites. A sampling depth in the sediment of e.g. <3cm ignores the many fine-grained sediments where the oxidized layer is only mm thick. So a sample that penetrates into the reduced layer of the sediment will over-state AVS effects on metals in the oxidized zones. And, oxidation state is not a permanent state in the upper layers. Mixing of sediments and resuspension create a complex micro- and macro- habitats with different AVS concentrations, changing continuously. The use of AVS assumes that the sulphides limit bioavailability universally, which may not be true, if animals oxidize sediments in their gut.

- Most of the experience of using SEM/AVS is from fresh water sediments. Bioturbation rates, quality and quantity of organic matter and redox regimes are very different from the marine environment. Although, conditions in some respects are different, the principles for lakes (and rivers) are the same as for marine areas, as regards: (i) spatial and temporal variation, (ii) resuspension, and (iii) vertical gradients of AVS in sediments.
- The stability of AVS-bound metals during resuspension and changes in the oxygen conditions due to bioturbation may be questioned. One might also wonder how important resuspension and sulphide oxidation are in polluted rivers in continental North Europe (cf. the Flanders database).
- Measurements of AVS may be operational dependent and inter comparisons of results from different labs may be difficult (i.e. handling of samples etc).
- Bioaccumulation of metals has been demonstrated from sediments where the metal should not have been bioavailable according to SEM/AVS. Theoretically it is reasonable to assume that most metals will be of reduced bioavailability when they are bound to sulphides because the stability constant of the product is so high. All data shows that sulphides proportionately reduce concentrations of metals in pore waters (Lee et al, 2000a). So, where experiments are designed to emphasize pore water as an exposure route, the effects of sulphides are quite evident. But, experiments emphasizing the pore water route does not completely reflect the conditions in nature (Lee et al., 2000 b, c).
- It has recently been shown in the RAR for Nickel (Draft January 2006) that the SEM/AVS does not work for benthos inhabiting the sediment- water interface, only for the truly deep sediment feeding *Lumbriculus*.

Finally, some comments related to the specific text in the document and some recommendations:

- A metal may be bioavailable without being toxic. At several places in the text these terms are used as synonymous terms.
- There are generally few data on SEM/AVS from Scandinavian sediment studies. Supplementary analyses are necessary.
- Direct measurements of bioavailability and bioaccumulation of metals in sediment living organisms in test arrangements (sediment aquaria with test organisms) have shown to be preferable. It would be of interest to compare SEM/AVS with direct measurements of bioavailability and bioaccumulation based on these tests.
- It is not specified if a standard default organic carbon value in EU of 5% is applicable both in freshwater sediments and marine sediments. It is considered far too high for marine sediments.
- The solubility products used for metal sulphides for ranking metals assume pure metal sulphide phases. This may not be entirely relevant in natural aquatic systems.
- Comment to paragraph 3.21. (p. 12, first line) where it is said that “for metals that do not bind with sulphides it may be worthwhile to explore if a relationship can be established between the observed toxicity levels and the presence of organic carbon”(and) “other ligands”. However, this seems not to be exclusively an issue for elements which do not (strongly?) bind to sulphides. Thus, e.g. Langston and others have shown strong linear relationships between Pb and As in soft parts of burrowing bivalves, and the corresponding Me/Fe ratios in (marine) sediment. So, either no AVS, or were Fe-oxides out competing these (cf. covariance between AVS and OC, mentioned above).
- Mercury and silver have the lowest solubility products, but are not mentioned in the sequential calculations of $\Delta\{AVS\}$ (p. 13 & 14).
- $\Delta\{AVS\}$ as well as total AVS is dependent of the other metals that bind to sulphides. Hence, also the concentrations of the metals that bind stronger to sulphides need to be known for the sites that are under evaluation.
- In the risk characterisation an AVS normalized PEC must be compared with an AVS normalised PNEC. This is done in Eq-18, but not in figure 9.

In addition recent studies show that the Me-AVS/SEM does not correlate with toxicity for a range of sediment organisms. In contrast the toxicity relates better to the overlying water metal content. It would be an alternative to use the AVS/SEM as a threshold below which no toxicity occurs (although this is also questionable as seen from above comments) and to make bioavailability corrections as preformed in soil and water.

Consequently, the SEM/AVS work has increased our understanding of some processes in sediments, but should not be presented as an exclusively used risk assessment tool.

Comments to “One- pager” no. 9.

This document is addressing the selection of an AVS default value which should be representative for the EU-region. The default value is based on the Flanders data set, which is the only major dataset on AVS which exists in Europe (200 samples).

A total AVS value of $<1 \mu\text{mol/g}$ dry wt. has been suggested as a default value for oxidised sediments (worst case scenario). This value based on the Flanders data set has been compared with a few samples from other countries, indicating that the default value is sufficiently conservative. It should be pointed out, however, that the number of samples compared with Flanders dataset is extremely low and that conclusions may not be justified. So, the advice would be that more data, from a Nordic perspective, should be collected to evaluate the proposed default value. This is in agreement with the conclusions in the document.

- Comments to Background document ch. 2.3.

There is an overall agreement that concentrations of total metals in sediments are not related to environmental risk and that estimates of the bioavailable fraction is a better approach. The AVS-approach assumes that the sediment is anoxic and that a reactive sulfide pool is present (iron monosulphides). Whether a sediment is anoxic or not depends on the level of organic matter in the sediments, rate of bioturbation etc. In marine sediments underlying well oxygenated water masses, the upper 5-10 cm of the sediment may be well mixed and oxygenated due to bioturbation. Hence, anoxic sediments appear at greater depths and the reactive pool of solid phase sulphide which binds metals predominates below the biological active surface layer. This should be pointed out in the background document.

In anoxic sediments the presence of ammonia would have a great influence on the toxicity to sediment living organisms and in some cases lead to misinterpretation with respect to metal toxicity. This should be mentioned in the background document.

It is stated in the background document that most of the studies reporting seasonal variability of AVS have been addressing uncontaminated sediments. More evidence should be presented on seasonal variations of AVS in contaminated sediments in a natural system (not spiked sediments).

It is pointed out in the document that bioturbation causes reduction of AVS in the surface sediments and remobilization of metals and increased levels of metals in the pore water. This indicates that AVS-bound metals may not be stable with respect to oxidation and resuspension, depending on the speciation of the metal sulphide formed. It may not necessarily be pure metal sulphides formed with a solubility product according to the text book, but rather a mixture of sulphides in an amorphous form.

It is stated in the document that metal in pore water will not necessarily be chemically available to benthic organisms, since any metal that is present in the pore water has the potential to form non-bioavailable metal complexes with other pore water ligands, thereby further reducing the potential for toxicity. This would imply that use of DGT-probes in sediments will not indirectly indicate bioavailability. This is an example how difficult it is to demonstrate bioavailability by a chemical approach. Generally, the MERAG document is not well balanced between chemistry and biology/ecology.

It is stated in the document that “since AVS levels decrease with increasing depth it is important to focus the analysis on the biological active layer of the sediment (0-20 cm). But even over this distance the AVS-profile could differ dramatically.” However, at least in lakes, most macro invertebrates are mostly found in the uppermost 5 centimetres and this depth should then be the most relevant when sampling for AVS.

References

- Dahlqvist, R., Benedetti, M.F., Andersson, K., Turner, D., Larsson, T., Stolpe, B. & Ingri, J. 2004. Association of calcium with colloidal particles and speciation of calcium in the Kalix and Amazon rivers. *Geochimica et Cosmochimica Acta* 68, 4059-5075.
- De Schampaelere, K. A. C & C. Janssen (2004), Development and field validation of a biotic ligand model predicting copper toxicity to *Daphnia magna*. *Environ. Toxicol. Chem.* Vol 23(6): 1365-1375.
- De Schampaelere, K. A. C & C. Janssen (2004), Effects of chronic dietary copper exposure on growth and reproduction of *Daphnia magna*. *Environ. Toxicol. Chem.* Vol 23(8): 2038-2047.
- Gustafsson, J.P. 2001. Modeling the acid-base properties and metal complexation of humic substances with the Stockholm Humic model. *Journal of Colloid and Interface Science* 244, 102-112.
- Koopal, L.K., Saito, T., Pinheiro, J.P. & van Riemsdijk, W.H. 2005. Ion binding to natural organic matter: General considerations and the NICA-Donnan model. *Colloids and Surfaces A265*, 40-54
- Lee B.-G., Griscom, S. B., Lee, J-S., Choi, H. J., Koh, C-H., Luoma, S.N., and Fisher, N. S. 2000a. Influence of dietary uptake and reactive sulfides on metal bioavailability from sediments. *Science*. v. 14, 287:282-284
- Lee, J-S., Lee, B-G., Luoma, S.N., Choi, H-J., Koh, C-H., and Brown, C. L. 2000b. Influence of acid volatile sulfides and metal concentrations on metal partitioning in contaminated sediments. *Environ. Sci. Technol.* 34, 4511-4516,
- Lee, B-G., Lee, J-S., Luoma, S.N., Choi, H-J., Koh, C-H. 2000c. Influence of acid volatile sulfide on the metal concentrations on metal

- bioavailability to marine invertebrates in contaminated sediments. *Environ. Sci. Technol.* 34, 4517-4523
- Tipping, E. 1994. WHAM – A chemical equilibrium model and computer code for waters, sediments and soils incorporating a discrete site/electrostatic model of ion binding by humic substances. *Computers and Geosciences* 20, 973-1023.
- Tipping, E. 1998. Humic ion-binding model VI: An improved description of the interactions of protons and metal ions with humic substances. *Aquatic Geochemistry* 4, 3-48.
- Unsworth, E. R., Wanken, K. W., Zhang, H., Davison, W., Black, F., Buffle, J., Cao, J., Cleven, R., Galceran, J., Gunkel, P., Kalis, E., Kistler, D., Van Leeuwen, H. P., Maryin, M., Noel, S., Nur, Y., Odzak, N., Puy, J., Van Riemsdijk, W., Sigg, L., Temminghoff, E., Tercier-Waeber, M-L., Toepperwien, S., Town, R. M., Weng, L. & H. Xue (2006), Model predictions of metal speciation in freshwaters compared to measurements by in situ techniques. *Environ. Sci. Technol.* Vol. 40: 1942-1949

ANNEX 6:

Fact Sheet 6 - “Incorporation of bioavailability for soils” and Chapter 2.4 in the background document “Bio-availability of metals/metals compounds in the terrestrial compartment”.

Reviewed by: *Janeck J. Scott-Fordsmand*,
National Environmental Research Institute (DMU), Denmark

Aim of the fact sheet

The fact sheet 6 and background report discuss the implementations of bioavailability into the ecological risk assessment of metals, especially focussing on soil.

Short description on the approach and proposed methods

Three main issues are introduced: 1) a translation of toxicity from total to pore-water or free-ion concentrations, 2) a soil-type-correction of toxicity data (regression based) and 3) a lab-to-field correction factor. In addition it is envisaged how this can be used in connection with the environmental risk assessment (ERAR).

Comments and suggestions

The implementation of the 3 issues is of trans-European nature. From a Nordic perspective this calls for careful analysis of the results as the underlying models and factors are based on central European conditions.

General

In general the suggested approach will enable a more refined risk assessment than previously possible, as soil-type models are introduced. The approach is solely bioavailability orientated (as suggested in the title), it is not discussing the drawbacks and ignores other important issues, such as mixture toxicity, competition and more ecological aspects. There is a lack of discussion of pros and cons of the applied approach; this is of course especially prominent for the background document. For example, what are the cons of using a free-ion approach? Are data available which

contradict that the lab-to-field factor is solely described by an ageing measure?

In the fact sheet words like “realistic worse case scenarios” are introduced, leaving no definition of what this may be. Further, a 10/90th percentile scenario is mentioned, where it should be discussed whether a 5/95th percentile is a better choice. This could be a serious discussion point in future ERARs.

From fig. 10 (fact sheet) it seems as all models should be employed, but clearly if the free-ion approach is used there is no need to perform a soil-type-correction on the effect side. The documents (especially the background document) have an unfinished appearance.

“Section 4.2 Use of transformation to pore water and use of speciation models”

For a “better normalisation” of the data it is proposed to advance through the data levels converting total concentrations to pore-water or free-ion concentrations. In principle the first attempted is to make the ERAR based on total metal concentrations. The second step is to test whether a more precise (not necessarily more accurate) ERAR can be based on pore-water concentrations or free-ion concentrations. The long-term agenda could be a terrestrial T-BLM.

For years it has been attempted to find a better estimate of exposure than the total concentration. This is because the total concentration in some case provides a poor fit with toxicity. Hence, it is excellent to explore whether pore-water or free-ion concentrations can be used when estimating environmental risk. The approach is based on more or less mechanistic models. It assumes a specific exposure route and that a point estimate of exposure reflects long-term toxicity.

Route of exposure

The approach to translate the total concentrations to pore-water or free-ion assumes that the pore-water is the main exposure route for organisms. This has not been discussed and needs to be verified for a wide range of organisms. For example, Bengtson et al (1983) found food ingestion to be the main route of exposure for *Collembola*. For earthworms the exposure route is not clear and may be compound dependent (Saxe et al 2001, Scott-Fordsmand et al 2004). Still, recent studies indicate that a terrestrial BLM may work for Cu in one earthworm species (Steenberger et al. 2005).

Lack of clarity and introduction of uncertainty

As it is difficult to measure pore-water and free-ion concentrations it is suggested to calculate this with various models. When more complex models are introduced there is a loss of clarity. It becomes unclear what uncertainty the models include, simply because the uncertainty surround-

ing each parameter is normally not included in the overall model. The magnitude of the cumulative uncertainty should be clearly shown and the model should be validated on an external data-set. The graphs in the background document seem to indicate only auto-validations. As alternatives to the advanced models it is suggested to use empirically derived K_d values. In this case it is also important to recognise the magnitude of the uncertainty related to such K_d estimates, and to validate the outcome. It is crucial to relate the magnitude of the uncertainty to the final outcome of the model. For example, if K_d the models are applied on the Nordic countries it should be discussed whether the uncertainty of the K_d estimate is larger than the natural span of K_d , and which implication this has for the risk assessment.

What do they represent?

These models are typically based on experiments (or soils) representing standard European conditions and not directly reflecting for example Nordic conditions. A good coverage of Nordic conditions with regard to the model parameters should be ensured.

Further, when proposing a risk assessment based on pore-water or free-ion concentrations, it should be considered that under field conditions the pore-water concentration and the free-ion concentrations will change continuously. It is recognised that this latter point has to some extent been discussed in the background document, but the subject calls for further discussion.

“Section 4.3 Use of toxicity related bioavailability models”

Normalisation of NOEC values per soil type: derivation of a soil sensitivity factor

A soil-type correction (normalisation) approach is suggested on the PNEC side by using regression models for the soil-parameters, e.g. CEC versus the toxicity estimates. Hence, after normalisation a NOEC/EC_x can be derived for a chosen standard soil.

In general this approach is interesting and relevant as clearly one single PNEC will not be able to cover all of Europe, not alone the Nordic countries. Nevertheless, some points need careful considerations.

A normalisation of the PNEC side would reflect the fact the normalisation equation in fact relates to a specific organism. This means that various organisms will be normalised with different equations. The normalisation should be done for each species, and the normalisation equations should be valid for that particular species. Furthermore, nothing is said about what should be applied on the PEC side?

The proposed approach is based on linear models, but it should be studied whether other models could be applied. Such models may include optimum or biphasic models. The model choice may be particular impor-

tant with regard to the soil-types on the edges of the original model. For example, if the true soil CEC – ECx relationship is best described by an optimum or a biphasic model, then a linear model may provide a worse prediction of low and high CEC soils than other models. The overall fit of the linear may still be adequate due to a large number of samples in the centre of the model. The models used in the risk assessment should employ soil-types that not only represent Europe, but also Nordic countries (from a Nordic perspective). This is equivalent to the soft water issues for the metals assessed within the existing substances regulation RARs. A small trans-Nordic project could be established analysing how Nordic soils perform in the current suggested trans-European models, e.g. Ni-models.

A crucial point is also for which species the models are derived and whether the models can be applied to other species. If the model is based on e.g. *F. candida* then it should be verified that this represents other arthropod species, before the model is applied to these. It seems unlikely that the model will also cover the predatory mite *H. aculeifer*. To minimize efforts this can be done by testing a range of preferably taxonomically different species at conditions representing the extreme of the *F. candida* model.

Finally, the underlying assumption for applying the model to other species is that the relative toxicity between species is fixed across soil type. Hence, can it be assumed that the model is the same for a second species as for the first species? This should be discussed and verified.

“Incorporating long term effects on metal bioavailability: derivation of a soil lab-to-field factor”

It is proposed to apply a lab-to-field factor to the toxicity data. The factor is based on a comparison between the toxicity for an organism when exposed in spiked soil and in aged/field contaminated soil.

The comparison between two laboratory tests, i.e. the test with the spiked soil and the test with the aged/field contaminated soil is not a laboratory-to-field extrapolation as such; it covers only a part of the extrapolation. In order to make a lab-to-field extrapolation additional issues should be considered. In the field organisms are usually exposed over generations and they are exposed to additional stressor such as climate, competitive stress and mixtures.

Several studies performed on a copper contaminated site, Hygum in Denmark (Bruus Pedersen et al 1999, Scott-Fordsmand et al. 2000 a,b, Strandberg et al 2006) highlight the laboratory-to-field extrapolation. In these studies, soil from this field site was brought into the laboratory and tests were performed with both newly spiked soil (soil from the site) and with field contaminated soil. Tests were performed with collembolans and earthworms. Exposing these organisms to Cu in the field contami-

nated soil indeed resulted in higher ECx values than seen from exposure in newly spiked soil. This indicates that “ageing” does occur. However, it was noted that the ECx values based on spiked soil corresponded better to the field threshold level (the threshold level was based on a multivariate analysis of the communities present along the field gradient) than the ECx values based on laboratory test with field contaminated soil. This observation strongly suggests that although “ageing” may take place, additional stressors have real importance in the field. Further, in the field the initial toxicity may have caused some populations to adapt, hence the threshold-level observed in the field represent adapted populations.

The lab-to-field extrapolation is suggested as a factorial approach. This approach should be made probabilistic rather than factorial. A probabilistic based approach will provide the risk assessor with an estimate of the uncertainty related to this value and will encourage data-suppliers to provide more data.

Much effort (and research) has been spent in refining the data input and the estimation of the HC5, e.g. ensuring an adequate number of species, good ECx-values, correction for soil type and more. The proposed lab-to-field factor which can have a large influence on final PNEC is based on only few results (not covering the span of different soil types, organisms and exposure situations). This is not advisable.

As it is proposed to introduce a lab-to-field factor based on “ageing”, it may also be proposed to include a factor that considers other lab-to-field issues. For example, to include a factor which considers that laboratory tests are “less than one generation” whereas the field case is always “multi-generation” exposures. To include a factor that considers combined stress in the field, mixtures, competition etc. A large number of issues advocating for an additional extrapolation factor could be proposed, but this is impractical and hence it is proposed to keep at simple factor in the risk assessment covering all. Still, it should be the goal to study and quantify the various differences between lab and field effects. In the long run this should be included in the risk assessment.

As only the effect of ageing on the bioavailability of metals is included in the “lab-to field factor” a more appropriate term for the correction factor (e.g. “ageing factor”) should be used.

Calculation of bioavailability Toxicity values

It is proposed to apply the lab-to-field factor to the toxicity data in order to obtain a corrected ECx. After this the background concentration (C_b) is added, then the normalising for differences between soils is performed. However, doesn't the lab-to-field factor already contain the background concentration, since an ECx obtained from an experiment with the contaminated soil must include the background concentration? Further, the soil sensitivity models are not based on lab-to-field corrected soils.

Tier 1 and Tier 2

Two approaches to including the correction factors are proposed, considering whether there are few or many data. If few data are available it is proposed to use the most conservative correction value (the Bio-F). This correction value should be based on the organisms for which soil sensitivity models are available. Given that models are only available for very few organisms this gives considerable uncertainty as to whether this is a conservative estimate. Further, relevant species (for regions) should be used.

This issue will not be discussed further as the main issue is to consider whether the correction factors should be used at all.

Conclusion

In conclusion the fact sheet and the background document have the appearances of an early draft version, with mistakes and lack of references for statements. The approach suggested can have their merits and may provide a tool for a more refined risk assessment, but there is an insufficient discussion of both the pros and cons of the various approaches. This lack of qualified discussion is especially prominent with the ageing approach.

References

- Bengtsson et al 1983. Growth changes caused by metal uptake in a population of *Onychiurus armatus* (Collembola) feeding on metal polluted fungi. *Oikos* 40: 216–225.
- Bruus Pedersen et al 1999. The impact of a copper gradient on a microarthropod field community. *Ecotoxicology* 8: 467–483.
- Saxe et al 2001. Novel model describing trace metal concentrations in earthworm, *Eisenia andrei*. *Environ. Sci. Technol.* 35: 4522–4529.
- Steenbergen et al 2005. Development of a biotic ligand model and a regression model predicting acute copper toxicity to the earthworm *Aporrectodea caliginosa*. *Environ. Sci. Technol.* 39: 5694–702.
- Scott-Fordsmand et al 2000a. Responses of *Folsomia fimetaria* (Collembola: isotomidae) to copper under different soil copper contamination histories, in relation to risk assessment. *Environ. Toxicol. Chem.* 19: 1297–1303.
- Scott-Fordsmand et al 2000b. The importance of contamination history for understanding the toxicity of copper to the earthworm *Eisenia fetida* (Oligochaeta: annelida) using the neutral-red retention assay. *Environ. Toxicol. Chem.* 19: 1774–1780.
- Strandberg et al 2006. Effects of a copper gradient on plant community structure. *Environ. Toxicol. Chem.* 25: 743–75

ANNEX 7:

Fact sheet 7: “Uncertainty analysis”, and Chapter 6 on the Background document: “Examples of probabilistic techniques and quantification of uncertainty in metal risk assessments”.

Reviewed by: *Tuomo Saloranta*
Norwegian Institute for Water Research (NIVA), Norway

Aim with the Fact sheet

The Fact sheet 7 gives a brief introduction to techniques and methods of uncertainty and sensitivity analysis, which can be used to cast light on and quantify the uncertainties in results from e.g. a risk assessment. Uncertainty analysis aims to replace point estimates with realistic distributions representing the uncertainty and/or variability in the parameters. Sensitivity analysis aims to find out how the different input factors or parameters affect the output variables, i.e. how sensitive the output is to the different input factors. These two important analysis techniques can enable the risk manager or decision maker to see more transparently, and in terms of probabilities, how realistic the different outcomes from the risk assessment are, and for which elements of the assessment the results are most sensitive.

Comments and suggestions

In the following, some more detailed comments on the three subsections of the Fact sheet 7 are presented.

Section 1: Introduction

- The distinction between uncertainty and risk should be made clearer (how is risk defined?)
- Figure 1 is very trivial and could be omitted.
- Figure 2. The decision tree implies that a probabilistic approach can be performed for substances that are not “data rich”, providing that prior knowledge on distribution shapes exists. One could argue that prior knowledge on distribution shapes must be based on substantial data, and therefore more data should be required in this case, before moving on to the node “Probabilistic approach”. Also, the pathway via nodes “deterministic approach” -> “risk?” -> “no further

refinement” is valid only when the deterministic approach is more conservative than the probabilistic approach, which one cannot conclude firmly on before both approaches are performed.

Section 2: UA for data-poor metals

- The terms “typical, average, realistic scenario” are somewhat confusingly used in the Fact sheet 7.
- Although the caption text of Table 1 states that the table represents an “indicative overview”, the information value and usefulness of this large two-page, still draft-like table for the reader may not be very high. Therefore, the whole Table 1 could be omitted. If it would be retained, its content should be revised (see bullet point below) and it should be emphasized that the table serves as an example of documentation of the selections made during such deterministic risk assessment, and thus serves as an important way to improve the transparency and reproducibility of the risk assessment.
- Table 1 contained also some unclear terms and some errors compared to the TGD (EC, 2003); e.g. under Step 5, “Derivation protection level”, which shall cover the variability due to inter-species sensitivity, HC5 would be the correct level in case a SSD is performed. For data poor substances the lowest NOEC or L(E)C50 value would be used. The assessment factor (10-1000) is applied in the derivation of PNEC (Step 6). For derivation of PNEC (Step 6) an application factor (1-5) is used only when it is based on a HC5 from a SSD. The table gives the impression that separate application factors are used to account for data quality, diversity, and representativeness, sampling uncertainty and lab to field extrapolation. In fact only one application factor (1-5) shall cover all these aspects according to the TGD.
- The “Weight of evidence approach” should be better clarified

Section 3: UA for data-rich metals

- Figure 7 shows uncommonly narrow uncertainty bands for the ECD-curve, and thus is not very representative as an example figure.

Conclusions

Generally, the Fact sheet 7 introduces quite well most of the cornerstones of uncertainty analysis (i.e. sources of uncertainty, sensitivity analysis, importance of transparency in the assessment, selection of probability distributions, correlations in the uncertainty analysis, Monte Carlo techniques) without getting into too many details. However, some revisions

are needed to improve clarity, and especially consistent use of different terms (now e.g. assessment factor and application factor are used as synonyms). Also a more rigorous definition of some key terms (such as “risk”) and consistent use of them in the text would be recommended. What is also lacking is a more thorough description of the different ways to categorize uncertainty (now only uncertainty vs. variability are discussed; see e.g. Funtowicz and Ravetz (1990) who made the distinction between technical, methodological, epistemological and ethical uncertainty), and more emphasis on the fact that 1) some types of uncertainty are very difficult to estimate quantitatively, and/or are irreducible, and that 2) therefore all uncertainty analyses will basically remain more or less uncertain. Especially the last point is very important to emphasize more (now it is only mentioned on the last few rows of the Fact sheet 7), i.e. to avoid giving the reader the impression that by uncertainty analysis one has covered and “overcome” all uncertainties in the results.

Although the text is theory-oriented, one should emphasize more the practical pitfalls in uncertainty- and sensitivity analysis (cf. background document’s chapter 6), e.g. the sensitivity to probability distribution’s tail values (methodological uncertainty), difficulties of estimating all parameter intercorrelations, difficulties in communication or presentation of the results from uncertainty analysis and the potential consequences of this for the stakeholder’s.

As pointed out above, the Fact sheet 7 gives a good introduction on the techniques of probabilistic assessment and the steps to be followed when performing an uncertainty analysis. However, the values presented in this Fact Sheet for practical application (e.g. in Table 1) are sometimes debatable, dubious, or incorrect. Thus it would be recommended that this Fact sheet 7 only focuses on the way to perform uncertainty- and sensitivity analysis on a general level, and does not go into practical details on what percentiles, application factors or other particular parameter values should be selected.

Comments to Chapter 6 in the Background document

(Examples of probabilistic techniques and quantification of uncertainty in metal risk assessments)

Chapter 6 in the Background document presents some specific themes as examples, and illustrates at least two important points: the value of the probabilistic approach, and higher sensitivity in the PDF’s tail. However, generally the text lacks depth and explanations and takes “shortcuts” as many examples are presented in a very brief manner. The text is also generally rather incomplete and draft-like (missing references, clutter in figure numbering & labels) and thus major revisions should be made to this chapter before it could be of any larger value to a wider audience.

Reference

- EC. 2003. Technical Guidance document on Risk Assessment *in support of* Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. European Commission, Joint Research Center, Institute for Health and Consumer Protection, European Chemicals Bureau, Ispra (VA), Italy.
- Funtowicz, S. O. and J. Ravetz, 1990: Uncertainty and quality in science for policy. Kluwer, Dordrecht, 229 pp.

ANNEX 8:

Fact sheet 8: “Data compilation, selection and derivation of ecotoxicity reference values. MERAG Program-Building Block: Classification for effects on the aquatic environment.” (Version June, 2006)

Reviewed by: *Jonas Falck* on behalf of the Nordic Classification and Labelling group.

Aim with the fact sheet

To develop a strategy on classification for effects on the aquatic environment of metals, metal compounds and alloys.

Short description on the approach and proposed methods

The approach taken is similar to and partly based on agreed strategies for environmental classification of metals and metal compounds (OECD 2001 and UN 2005), e.g. a comparison of transformation/dissolution and toxicity reference values.

However, methods and approaches globally agreed as guidance for classification purposes have been mixed with new, interesting but currently not applicable approaches; approaches that are sometimes contradictory to agreed strategies, e.g. the use of the Critical Surface Area (CSA) approach in metal classification.

Comments and suggestions

General

”*Guidance on the use of the harmonised system for the classification of chemicals which are hazardous for the aquatic environment*” have been agreed among regulators in collaboration with Industry, published by the OECD (OECD 2001) and incorporated into the Globally Harmonised System for Classification and Labelling of Chemicals (UN 2003 and UN 2005).

These guidance (including chapter 7 “Classification of metals and metals compounds”) is fundamental in environmental hazard classification. The agreed content of this document should not be changed by the development of alternative guidance. Instead, in MERAG:

1. the guidance referred to above should be reproduced just as they have been agreed,
2. parts that need further guidance should be identified, and
3. additional guidance could be proposed. This, however, should preferably be done within the frame of existing guidance on GHS. This is especially important if the Fact Sheet is meant to work as guidance to metal industry on environmental classification of metals.

In the light of the statement above the following comments on Fact Sheet 8 are made:

1. The structure of the text makes the agreed guidance (including strategies for environmental classification of metals and metal compounds) and new concepts (on e.g. use of BLM in derivation of reference value) difficult to distinguish from each other.
2. This document is difficult to use as guidance when already agreed and used strategies are mixed with new, interesting but currently not applicable approaches. Especially since some of the new concepts are clearly in contrast with the frame of agreed guidance on classification of metals and metal compounds, i.e.:
 - a) Derivation of reference value (grouping of data). There is already guidance on how to use aquatic toxicity data in derivation of reference value for classification purposes (i.e. A.9.7.2. & A.9.3.4, UN 2005). If there is a need for a clearer guidance, it should be based on that frame.
 - b) Use of the Critical Surface Area (CSA) approach in metal classification: It is clearly stated in already agreed existing guidance that “while this approach is not normally used for classification it may provide useful information for labelling and downstream decisions” (A9.7.5.4.4, UN 2005). How it could be used *for labelling purposes and downstream decisions* could, however, be a subject for development of further guidance.
3. Classification of alloys. Guidance for environmental classification of alloys is clearly an area that needs still to be developed. The approach presented in Annex to the Fact Sheet 8 is therefore welcomed as a draft it should be subject for further discussion by relevant bodies.

Conclusion

Since globally agreed guidance for classification purposes on aquatic hazard of metals and metal compounds already exist, there is no need to include this into MERAG. Hence, fact sheet No 8 should be replaced with references to agreed guidance (OECD 2001 and UN 2005).

Deficiencies in agreed guidance, if any, could be identified, but suggestions and proposals for their improvements should be brought forward to relevant bodies (i.e.OECD/UN and RIP3.6) for further development and agreement.

Reference

(OECD 2001). Guidance Document on the use of the Harmonised System for the Classification of Chemicals which are Hazardous for the Aquatic Environment. OECD Environmental Health and Safety Publications, Series on Testing and Assessment No. 27 Organisation for Economic Co-operation and Development, March 2001.

(UN 2003). Globally Harmonized System for Classification and Labelling (GHS). ST/SG/AC.10/30. United Nations, New York and Geneva, 2003. (UN 2005). Globally Harmonized System for Classification and Labelling (GHS). ST/SG/AC.10/30/Rev. 1. United Nations, New York and Geneva, 2005.

Sammanfattning

Syftet med föreliggande rapport har varit att kritiskt granska ett förslag på principer för riskbedömning av metaller (MERAG - Metals Environmental Risk Assessment Guidance), vilket tagits fram av industrin i samarbete med engelska miljöministeriet, under 2005-2006.

Projektet initierades för att förbereda för en diskussion om MERAG-dokumenterna i arbetet med s.k. Existerande Ämnen (Existing Substances Regulation, EEC 793/93), samt ett eventuellt kommande förslag att implementera MERAG i de vägledningsdokument som tas fram för den nya kemikalielagstiftningen (REACH) ,

MERAG utgörs av 8 faktablad, som behandlar olika aspekter av riskbedömning, samt ett bakgrundsdocument med motiveringar och förklaringar till de olika förslagen i faktabladen

Som underlag för föreliggande rapport har de olika faktabladen granskats av olika Nordiska konsulter och deras utvärderingar finns vidlagda denna rapport som fristående bilagor (1-8). Projektets styrgrupp (NO-RAP) har sammanfattat slutsatserna från de olika utvärderingarna i kapitel 3 av denna rapport.⁷

Förutom förevarande rapport, har detta projekt även förhoppningsvis inneburit en ökad insikt och delaktighet, bland Nordiska forskare, i principerna för miljöriskbedömning av kemikalier inom vissa EU-lagstiftningar

Det är tydligt att speciell riskbedömnings metodik behöver utvecklas för oorganiska ämnen, som metaller. Bland annat för att man måste ta hänsyn till naturlig förekomst, att vissa metaller är essentiella samt den historiska användningen av metaller. För vissa metaller är det nödvändigt att ta hänsyn till dess biotillgänglighet, för att kunna göra användbara riskbedömningar. MERAG dokumenten utgör därför en god början till utveckling av riskbedömningsmetodik som tar hänsyn till dessa faktorer.

Vid ett gemensamt möte (Oslo, februari 2006) med de inblandade konsulterna framkom ett antal gemensamma slutsatser om innehållet i MERAG. För det första ansågs inte Nordiska förhållanden vara tillräckligt täckta i de olika tillgängliga modellerna som används för att korrigera för biotillgänglighet. För de metaller där man hittills använt sig av biotillgänglighetskorrigeringar (Cu, Zn, Ni, samt delvis Cd och Pb), har databaser som inte är representativa för EU, använts för att ta fram "typiska" och "värsta fallet" scenarier när det gäller abiotiska faktorer, samt för att sätta gränserna för biotillgänglighetsmodeller. För sötvatten föreslås istäl-

⁷ Slutsatserna utgör inte med nödvändighet formella ståndpunkter för de olika Nordiska myndigheterna.

let att FOREGS databasen används för dessa ändamål. Metallkoncentrationer från denna databas kan, med försiktighet, användas som mått på dagens bakgrundskoncentrationer.

För det andra, måste bakgrundsdokumentet till MERAG revideras och den vetenskapliga kvalitén förbättras.

En tredje gemensam slutsats från Oslo-mötet var att MERAG dokumenten är obalanserade när det gäller kriterier för selektion av data. Mycket mer vikt har lagts vid att diskutera kriterier för acceptabla data från toxicitetsstudier, jämfört med kriterier för acceptabla exponeringsdata⁸.

Granskaren av faktablad No 1 stöder den generella riskbedömningsprocedur som föreslås i detta faktablad. Denna innebär ett stegvis förfarande, men ett villkor är att det första steget baseras på konservativa uppskattningar av PEC och PNEC. Vid risk för effekter i den initiala analysen, förfinas riskbedömningen t.ex. genom användande av biotillgänglig fraktion (istället för totalhalter eller löst fraktion av metallen) vid beräkning av PEC och PNEC. Begränsningar i använda biotillgänglighetsmodeller måste dock klargöras. Behovet av en extra osäkerhetsfaktor (minst 2) vid framtagande av PNEC utifrån 5:e percentilen (HC₅) av testorganismernas känslighetsfördelning (SSD) understryks.

I faktablad No 2 om exponering fokuserades granskningen på avsnitten om exponeringsanalys med hjälp av monitoringdata. Granskaren konstaterar att det saknas vägledning/kriterier för att bedöma kvalitet och representativitet i monitoringdata. Vidare, används olika termer för bakgrundshalter vilka bör definieras. Vägledning bör ges angående betydelsen av säsongvariationer vid analys/val av monitoringdata. Enligt TGD antas 90:e percentilen av monitoringdata representera "värsta fallet"-scenario. Granskaren anser att det vore mer logiskt att istället använda 95:e percentilen för detta ändamål, eftersom detta värde sedan jämförs med 5:e percentilen av effektdata.

Granskaren av faktablad No 3, om effektbedömning, konstaterar att diskussionen om acklimatisering/adaption är dåligt underbyggd. Detta är speciellt viktigt eftersom ett nytt koncept om "metalloregion" introduceras, där man föreslår att organismer har olika känslighet för metaller i olika geografiska områden och att man vid riskbedömning skall ta hänsyn till detta. Vidare kommenteras även av denna granskare att en extra osäkerhetsfaktor (>2) måste appliceras på HC₅-värdet, vid beräkning av PNEC utifrån SSD. Detta eftersom det finns osäkerheter som inte täcks av SSD, t.ex. ytterligare exponeringsvägar, andra stressfaktorer, betydelsen av nyckelarter och processer på ekosystemnivå. Förslagen att minska antalet arter som bör ingå i en SSD för sediment och markorganismer, jämfört med SSD för pelagiska arter, måste underbyggas bättre.

Granskaren av faktablad No 4, om marin risk bedömning, är mycket kritisk till beskrivningen av det marina systemet och förslagen som ges i faktabladet angående t.ex. extrapolering av toxdata för sötvattenarter till

⁸ NORAP-gruppen inser dock att detta inte är någon fråga som är specifik för metaller.

det marina ekosystemet. Detta innebär att han också är kritisk till den generella riskbedömningsmetodik som beskrivs i TGD för marina system.

Faktablad No 5 behandlar biotillgänglighet av metaller i vatten och sediment. Granskarna anser att BLM konceptet, som är den metod som förs fram för att modellera biotillgängliga former av metallen i vatten, är bra och bygger på ny forskning. Det finns dock en tendens i MERAG-dokumenterna till undervärdering av osäkerheterna i BLM-modelleringen, t.ex. extrapoleringen från ett par arter för vilka BLM-modeller tagits fram till att gälla för hela ekosystemet. Detta måste diskuteras utförligare.

För biotillgänglighetskorrektioner i sediment fokuseras vägledningen i MERAG på bindning av metaller till sulfider genom applicering av den s.k. SEM/AVS metoden. Granskarna är eniga om att effekter i sediment oftast inte kan relateras till totalhalter av metaller i sediment, men trycker på att SEM/AVS metoden är förknippad med mycket stora osäkerhet. Exempel på osäkerheter är relevansen av metoden för bentiska organismer med olika födoval och gräv beteenden, liksom betydelsen av variationen av AVS pga säsongvariationer i abiotiska förhållanden och variationer i recipientens morfologi och hydrologi. Det finns alternativa metoder för att relatera toxiciteten till vissa förekomstformer av metallen i sediment som också bör diskuteras. Exempel på sådana metoder är att relatera toxiciteten till halterna i den vattenfas som organismerna exponeras för (porvattnet eller vid sedimentytan) eller att göra biotillgänglighetsmodeller liknande de som föreslås för vatten (BLM) och jord.

I faktablad 6 ges förslag på jord-typ baserade modeller för implementering av biotillgänglighet vid riskbedömning av metaller i mark. Granskaren anser att förslagen innebär en förfining av riskbedömningsmetodiken, men trycker på att även andra aspekter än ren biotillgänglighet bör diskuteras, t.ex. om modellerna är applicerbara på alla arter. MERAG föreslår en omvandlingsfaktor från lab-till-fält för att ta hänsyn till 'åldring' (inbindningen av metallen i marpartiklarna). I denna bör även andra skillnader vägas in, såsom ekologiska och klimatrelaterade stressfaktorer. Det finns en tendens i MERAG-förslagen att endast inkludera faktorer som medför reduktion av risken.

Faktablad No 7 som behandlar osäkerhetsanalys, ger enligt granskaren en god introduktion i osäkerhetsanalysens grunder. Vissa nyckeltermerna bör dock definieras mer stringent. Det bör också betonas starkare att vissa typer av osäkerhet är väldigt svåra att kvantifiera och/eller inte möjliga att reducera, och att all osäkerhetsuppskattning därför alltid är mer eller mindre osäker.

Granskarna av faktablad 8, som rör klassificering av metaller, anser att det inte finns något behov av ny vägledning för klassificering av metaller och metallföreningar, eftersom det redan finns globalt överenskommen vägledning. Sålunda, bör MERAG faktablad 8 bytas ut mot den överenskomna vägledningen.

Abbreviations and acronyms

| | |
|--------------------|--|
| AVS | Acid Volatile Sulphide |
| BLM | Biotic Ligand Model |
| CEC | Cationic Exchange Capacity |
| CSA | Critical Surface Area |
| DMU | Danmarks Miljøundersøgelser |
| DOC | Dissolved Organic Carbon |
| EC | European Commission |
| ECD | Exposure Concentration Distribution |
| EC _x | Effect Concentration (x=% effect level) |
| ERAR | Environmental Risk Assessment Report |
| ESR | Existing Substances Regulation |
| EUROMETAUX | European Association of Metals |
| FIA(M) | Free Ion Activity (Model) |
| FOREGS | Forum of European Geological Surveys |
| GHS | Globally Harmonised System (for classification) |
| HC ₅ | Hazard Concentration, 5 th percentile |
| ICMM | International Council on Mining & Metals |
| ITM | Institutionen för Tillämpad Miljövetenskap |
| L(E)C _x | Lethal or Effect Concentration (x=% effect level) |
| LOEC | Lowest Observed Effect Concentration |
| Me | Metal |
| NIVA | Norsk Institutt for Vannforskning |
| NOEC | No Observed Effect Concentration |
| NORAP | Nordic Risk Assessment Project |
| OC | Organic Carbon |
| OCEE | Optimal Concentration range for Essential Elements |
| OECD | Organisation for Economic Co-operation and Development |
| PEC | Predicted (environmental) Exposure Concentration |
| PNEC | Predicted No Effect Concentration |
| RAR | Risk Assessment Report |
| REACH | Registration Evaluation Authorisation of Chemicals |
| RIP | REACH Implementation Projects |
| SSD | Species Sensitivity Distribution |
| TGD | Technical Guidance Document |
| UN | United Nations |
| UK | United Kingdom |
| VRAR | Voluntary Risk Assessment Report |
| WHAM | Windermere Humic Aqueous Model |