Fraunhofer

IME

FRAUNHOFER-INSTITUT FÜR
MOLEKULARBIOLOGIE UND ANGEWANDTE OEKOLOGIE IME

Christoph Schäfers

Ecological approaches to aquatic ecotoxicology challenged by the needs of risk assessment

FRAUNHOFER VERLAG

Fraunhofer-Institut für
Molekularbiologie und Angewandte Oekologie IME

Christoph Schäfers

Ecological approaches to aquatic ecotoxicology challenged by the needs of risk assessment

FRAUNHOFER VERLAG

Contact:
Fraunhofer-Institut für Molekularbiologie und Angewandte Oekologie IME
Auf dem Aberg 1
57392 Schmallenberg, Germany
Phone +49 2972 302-270
Fax +49 2972 302-319
URL www.ime.fraunhofer.de

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliografic data is available in the Internet at <http://dnb.d-nb.de>.
ISBN: 978-3-8396-0542-4

Lan 1

Zugl.: Koblenz-Landau, Univ., Habil., 2012

Printing and Bindery:
Mediendienstleistungen des
Fraunhofer-Informationszentrum Raum und Bau IRB, Stuttgart

Printed on acid-free and chlorine-free bleached paper.

Fraunhofer Information-Centre for Regional Planning and Building Construction IRB
P.O. Box 80 04 69, 70504 Stuttgart, Germany
Nobelstrasse 12, 70569 Stuttgart, Germany
Phone +49 711 970-2500
Fax +49 711 970-2508
E-Mail verlag@fraunhofer.de
URL http://verlag.fraunhofer.de

Christoph Schäfers

Ecological approaches to aquatic ecotoxicology challenged by the needs of risk assessment

Habilitationsschrift

zur Erlangung der Venia legendi

für das Fachgebiet Ökotoxikologie

vorgelegt dem Fachbereich 7:

Natur- und Umweltwissenschaften

der Universität Koblenz-Landau

Campus Landau

Landau, den 11.1.2012

Table of contents

Table of tables

Table of figures

Abbreviations

2-GT	Two-Generation Test
11-kT	11-keto-testosterone (active male androgen in fish)
21d-FA	Fish screening assay, OECD-validiert: 21d Fish Assay (OECD TG 230)
BVL	Bundesamt für Verbraucherschutz und Lebensmittelsicherheit (*German federal agency for consumer protection and food safety*)
EC50, EC10	Concentration associated with an effect level of 50, 10% compared to untreated controls, calculated by regression statistics
EDSTAC	US Endocrine Disruptor Screening and Testing Advisory Committee
FLCT	Full Life Cycle Test
FSDT	Fish Sexual Development Test (OECD TG 234)
FSTRA	Fish screening assay, US-validiert: Fish Short Term Reproduction Assay (OECD 229) (=21d-FA + fecundity + gonad histology + secondary sex characteristics)
GLP	Good labotarory Practice; international legal requirement for the performance, quality assurance and documentation of regulatory studies; to be certified
HC5	Hazardous Concentration for 5% of the species; percentile calculated from an SSD by regression statistics
IVA	Industrieverband Agrar (*German agrochemical producers association*)
JKI	Julius Kühnen-Institut (*German federal agricultural research agency*)
LOEC	Lowest observed effect concentration: lowest test concentration causing a significant deviation from the control, derived by statistical hypothesis testing
MoA	Mode of Action
MT	Multigeneration Test (e.g. MMT: Medaka Multigeneration Test)
NOEC	No observed effect concentration: highest test concentration without significant deviation from the control
NOEAEC	No observed ecologically adverse effect concentration: highest concentration without adverse effect; slight and temporary effects with fast recovery are not regarded adverse
PEC	Predicted environmental concentration: exposure concentration calculated by assuming conditions defined by agreed scenarios
PLCT	Partial Life Cycle Test
PRC	Principal response curves: Output of redundancy analysis, a multivariate statistcal evaluation method related to the principle component analyses with time as covariable
RAC	Regulatory acceptable concentration
RWTH	Rheinisch-Westfälische Technische Hochschule (*Technical University of Aachen*)
SD	Standard deviation from the arithmetic mean, calculated from the variation of representative samples (n-1)
SSD	Species sensitivity distribution: Figure of cumulative distribution of species sensitivities to defined statistical endpoints
TER	Toxicity: exposure ratio: Relation of a toxicity endpoint (e.g. NOEC for long-term tests, LC50 for acute tests) and the PEC
UBA	Umweltbundesamt (*German federal environment agency*)
VTG	Vitellogenin (yolk precursor protein)
WFD	EU Water Framework Directive

Preface

Scope of the thesis

Aquatic ecotoxicology is a field of research and regulation with contradictory objectives, items and approaches. It comprises different protection aims and philosophies to reach them. During my work as teacher and trainer, researcher, contract study director and consultant for regulatory authorities, I had the opportunity to approach the field from different views and focusing on different aspects. It is the aim of this thesis to give a consistent overview upon the field of experimental and evaluative aquatic ecotoxicology based on the own work, including some general conclusions.

Way of referencing

Because of the summarizing character, only own publications are completely referred to by using reference numbers. The order follows the year of publication in the publication categories: theses, peer reviewed papers and articles, monographs, invited papers, non-GLP reports not covered by publications, and selected GLP higher tier study reports of SSDs, micro-/mesocosm studies, extended fish studies in microcosms, fish full life cycle studies.

The other publications are referenced to by author and year. These are mainly publications in the context of regulatory and scientific concepts, or specific references not mentioned in the own publications, e.g., because reviews have not been published yet or discussions have gone further after publication.

For full references, please be kindly asked to make use of the own publications referred to by numbers.

Structure

The different chapters present specific approaches and methods of aquatic ecotoxicology. At the beginning of each chapter, I try to include a short assessment of the importance of the own contribution / key publications to the advance of research in the respective field.

Dedicated to Prof. Dr. Alfred Seitz

✝ January 15, 2010

Executive Summary

Ecotoxicology is a field of research and regulation with contradictory objectives, items and approaches. Its overall objective of ecotoxicology may be defined as **Identification and / or assessment of ecological effects of chemicals, minimizing uncertainties of cause and risk.**

Aquatic ecotoxicology. Biochemical reactions as basis of life performances need the aqueous solution and are characterized by polar and hydrophobic interactions. Exchange processes of toxicologically active substances between water as environmental medium and the aqueous solution associated with lipophilic compartments in organisms are the most effective ones. Consequently, aquatic organisms with high exchange rates at outer membranes exhibit the highest hazard potential, as long as toxicologically sensitive molecular structures are present. Aquatic ecotoxicology should identify or asses effects of chemicals on aquatic ecosystems. The focus of research can differ according to the objective and challenged by different needs. In a regulatory context, either a comprehensive evaluation of a substance is necessary for registration, notification or use authorization, or water quality has to be evaluated depending on type and use of the water body. For both regulatory objectives, science-based objectives have to be met, e.g. demonstrating physiological or taxonomic representation of the test organisms in terms of exposure pathway and toxicological sensitivity, or investigating and considering causal relationships of effect generation, manifestation and compensation on different levels of biological organization, i.e. biomarkers, cells, individuals, populations or communities. Finally, the technical objective is to develop, improve and standardize methodologies to meet the regulatory objectives by being as much scientifically appropriate as possible.

Substance evaluation

Regulatory needs. Substance-specific aquatic ecotoxicity data represent a substantial part of information needed for registration and labeling of chemicals and products, triggering various further regulations. The data set should be protective for all aquatic systems at risk (precautionary principle), but it should not be too overprotective when overruling economic interests of producers and consumers. For these reasons, the assessment should be oriented at the realistic worst case. As a basic step, it is necessary to assess the aquatic hazard potential of a chemical by determining its intrinsic aquatic toxicity. In a second step, the risk for aquatic systems may be assessed by comparing the intrinsic toxicity with estimations of exposure, which considerably differs between the regulatory frameworks for pesticides, industrial chemicals, biocides, veterinary and human pharmaceuticals.

Due to the limited resources of time, money and expert knowledge, it is not useful to test all chemical substances in depth. Thus, in all legislative contexts tiered approaches have been implemented starting with similar tests at the lower tiers, which should well represent aquatic systems to detect a maximum of potential effects at lowest costs. As the representation by the toxicity tests is always limited, an application factor to the results has to be applied to address the uncertainties of extrapolation from the basic test set data to safe concentrations for aquatic communities. These are uncertainties about more sensitive species, uncertainties about potentially more sensitive life stages and performances (chronic effects), and uncertainties about secondary effects. These uncertainties become critical with increasing exposure potential and can be reduced by performing tests with additional standard test species and standard chronic tests. If the environmental risk assessment indicates missing safety margins between predicted environmental concentration and precautionary toxicity estimates, higher tier test can help to further reduce uncertainty of the risk assessment and enhance the justification to expose aquatic ecosystems to high concentrations, as adverse effects can be excluded with high probability. For this purpose, the specific concerns of the chemical product have to be identified and addressed in the studies, which should include ecological realism as well as representation of sensitive species, life stages and worst case exposure conditions. The preparation of complex test systems with sensitive endpoints and high statistical power/low variability and their maintenance for the time of investigations needs ecological understanding, experience and technical expertise.

Higher tier risk assessment study approaches. The described ecological approaches to aquatic ecotoxicology comprise EC projects, higher tier studies for industry, R & D projects for authorities and own projects supported by Fraunhofer. The main topics are 1) the representative investigation and interpretation of especially variable invertebrate species sensitivities, 2) the consideration of sensitive life stages and performances in the development of full life cycle tests with fish and the derivation of a fish test strategy for endocrine disrupting chemicals, and 3) the development of an indoor semi-realistic microcosm system for the representative investigation of community effects regarding specific concerns and exposure patterns.

1) The presented projects on **invertebrate species sensitivity** focus on physiological traits depending on taxonomical and ecological relations, e.g. the adaptation to the conditions of the specific habitats. A comparison of species sensitivities of groundwater and surface water organisms to selected pesticides revealed that inhibitors of anabolic pathways or neuronal activity affect groundwater organisms less or clearly retarded as compared to surface water organisms. No hints to a physiologically higher sensitivity of groundwater organisms could be observed. Thus, the sensitivity distribution of lethal toxicity to groundwater species can be conservatively assessed by testing taxonomically and physiologically comparable species of surface waters. However, it has to be strongly emphasized that the population dynamics of groundwater species, if affected, allows only very slow and insufficient recovery by reproduction or recolonization.

   For surface water organisms and ubiquitous modes of action (acetylcholine-esterase inhibition as shown with carbaryl), species sensitivity seems to be mainly due to uptake and internal exposure, depending on the water exchange probability at respiratory membranes and thus being highly correlated with oxygen demand. This is clearly habitat specific. The presented examples demonstrate that including ecological information can substantially reduce uncertainty. For meeting the requirements of higher tier ecological risk assessment, SSDs should be used in a two-step-approach: a) include species of all relevant taxonomic groups to ensure the identification of sensitive groups. All information on exposure and hazards should be considered when selecting test species. The slope of the distribution will be flat, the fit may not be satisfactory, the HC5 will be very low and uncertain, and thus of limited value. b) cluster species according to taxonomy / physiology and / or habitat ecology. Some clusters may be excluded from further assessment due to obvious insensitivity or missing exposure. The risk assessment for the remaining sensitive clusters can either be performed separately (including further approaches like realistic exposure), or by calculating an SSD of combined species. The slope will be steeper, the HC5 will be higher, and the uncertainty of the assessment will be lower. Nevertheless, lowest toxicity values near or even below the HC5 should be evaluated for the ecological severity of statistical endpoint. Representative chronic risks should be accounted for. If the SSD is based on EC50s instead of NOECs or EC10s, habitat exposure and recovery potential should be considered. This will reduce substantial uncertainty of the overall assessment beside the reduction of statistical uncertainty.

2) The presented projects on **sensitive life stages and performances in fish** focus on sexual endocrine effects, population relevance and specificities of different regulations. During 24 fish full life cycle tests (FLCT), two- or multi-generation tests (2GT, MT) in public or confidential GLP projects, of which 17 tests investigated endocrine disruption by different mechanisms of action, we completed knowledge on mechanism-specific

responses, derived endpoint-specific quality criteria and developed test protocols and a test strategy for the indication, hazard and risk assessment of endocrine effects.

For definitive testing of potential endocrine disruptors, population relevant or apical endpoints were opposed to endpoints indicative for specific modes or mechanisms of action. Population relevant endpoints are hatch, stage-specific survival, growth, time to first spawning (= duration of sexual development), sex ratio, fecundity and fertility. Indicative endpoints such as secondary sex characteristics, histopathology, physiological biomarkers or gonado-somatic index are not necessary for risk assessment purposes, but can be important to understand and interpret apical effects and attribute them to modes of action for hazard-based regulatory purposes.

Especially but not only for pesticide regulation, effects of peak exposure have to be assessed. Long-term fish test protocols were developed for intrinsic hazard assessment. They are based on constant exposure. We developed a zebrafish full life cycle test design for worst case peak exposure in a static water/sediment system with artificial sandy sediment. Three different fish life stages (embryos, juveniles and adults) are concurrently exposed to a single or repeated application of the test substance. The unique study type investigates the potential of the test substance and its metabolites generated in the water or sediment for prolonged fish toxicity, early life stage toxicity, impact on juvenile growth, impact on reproduction, and early life stage toxicity of the filial generation, following peak exposure to all three introduced life stages. As the three life groups per vessel run through the same life stage at different time of exposure and consequently at different test substance concentrations, multiple comparisons of effects are possible within and between replicates. The relation of effects to time-weighted average concentrations during susceptible life stages can be compared to or used as intrinsic toxicity data. Endpoints can be differentiated according to their recovery potential and thus permanence and severity of effects. Thus, in zebrafish sex reversal from females to males did not show recovery. Retardation of growth and male development including fertilization capacity recovered without exposure and under presence of appropriate feeding and competition conditions in weeks to months, depending on severity of the primary effect. Fecundity recovered in days, as in the performed studies it was mainly dependent on spawning activities, which started again soon after exposure decreased.

Regarding sexual-endocrine disruption, the developing intrinsic test protocols for short-term in vivo-screening (FSTRA, 21d-FA) and long-term tests (FSDT, FLCT, 2-GT and MT) were used for the identification of effects of estrogens, anti-estrogens, an androgen and an anti-androgen, aromatase inhibitors and other inhibitors of steroid synthesis. The results and discussions with the German UBA resulted in important regulatory considerations: A testing strategy consisting of a basic tier identifying the potential of sexual endocrine disruption (structural information and read-across, *in vitro*-screening data, information from toxicological of efficacy studies), a tier 1 (*in vivo*-Screening, 21d-FA) identifying endocrine disruption at relevant concentrations in water and a tier 2 definitive test for population-relevant hazard assessment is sufficient. IME contributions are:

- *Basic tier***.** If the acute: chronic ratio (ACR) in fish is high, and if no other MoA is obvious, concern of endocrine disruption cannot be excluded. From experience with usual ACRs by industry and research, an ACR>20 is proposed as trigger value.
- *Tier 1.* As in a tiered testing and assessment approach the possibility of false negative findings should be minimized, the primary strategy by EDSTAC to implement three *in vivo* tiers (fish screening assay, short-term reproduction/partial life cycle test and two-generation test) is not appropriate. Either the tier 2 test (short-term reproduction/partial life cycle test) covers the most sensitive exposure and effect manifestation phases and can replace the definitive test, or not. In the

latter case, it will lead to a false negative response. Thus, the FSDT (OECD 234) can be used as definitive test provided that the most sensitive exposure window and the most sensitive manifestation of effects are covered. To date, this can be assumed for aromatase inhibition and androgen receptor agonists. The short-term reproduction assay or partial life cycle test starting with reproduction and ending with the F1 early life stages is not suited as definitive test, as all sexual endocrine mechanisms of action were shown to exhibit the sexual development phase as most sensitive exposure window, even when effects become manifest at later stages (reproduction, filial early life stages).

Regarding the predictive potential of the biomarker responses identified by IME zebrafish test, VTG increase is indicative of estrogen receptor agonists, VTG decrease in females is indicative of androgenizing effects by estrogen receptor antagonists and aromatase inhibitors, provided that there is not a complete sex reversal. 11-kT increase indicates compensatory reaction to androgen receptor antagonists. VTG and 11-kT decrease may indicate a general steroid synthesis inhibition, either by disruption of testosterone synthesis or by systemic toxicity. Sensitivities in short-term tests (21d-FA, OECD 230) are within the same order of magnitude as population-relevant effects observed in definitive tests. Androgen receptor agonist is the only mechanism shown to deviate, as biomarker responses are by far less sensitive than sex reversal. In case of potential androgen effects a test species with easy observable sex and secondary sex characteristics that are influenced in the first three weeks of exposure should be used, at least until an appropriate biomarker has been found.

- *Tier 2.* For a definitive test we propose a 2-GT starting with fertilized eggs and ending after sexual maturation of the F1-generation. Compared to the medaka MT (MMT) put forward by the Japan and the US, this design ensures the most crucial exposure phase during sexual maturation in the two generations and thus maximum possibilities of maternal transfer of effects. At the same time it reduces fish numbers by one quarter, test duration by 17% and omits the most variable reproductive phase, which has already been tested in the screening assays. Nevertheless, IME is developing a zebrafish MT according to the MMT.

3) The value, possibilities and limitations of **community level testing** approaches for regulatory risk assessments were discussed at the CLASSIC workshop in Schmallenberg organized by the author in 1999 and further specified at the AMPERE workshop in Leipzig 2007. Main and partly contradictory aspects of community level studies are a) realism concerning exposure and communities of agricultural landscapes, b) inclusion of representatives of the most sensitive species known from lower tier testing or literature, c) extrapolation possibility to a wide range of natural conditions. Uncertainty factors applied to the NOEC or NOEAEC have to account for realistic worst case conditions concerning exposure (partitioning and dissipation in the system) and sensitivity of endpoints (community composition, abundance of sensitive species, functional endpoints).

   IME developed 1 m³ indoor semi-realistic microcosm systems with sunlight simulation and temperatures control between 5 and 40 ± 1°C for the representative investigation of community effects regarding specific concerns and exposure patterns and to reduce uncertainty of assessments based on outdoor mesocosm studies concerning extrapolation to different conditions.

   Pesticide registration studies. Examples of studies performed in the IME systems addressing specific concerns of herbicides, insecticides/acaricides and fungicides are presented and discussed in comparison to outdoor approaches. The smaller volume and the indoor situation is a clear disadvantage with respect to biodiversity. Dominance patterns are more pronounced. Abundances of larger macroinvertebrate or macrophyte species tend to be too low for statistical evaluation. These communities should be investigated in larger outdoor systems. However, manageability and constancy of conditions are clearly better in the smaller indoor systems.

- Thus, there are advantages to test substances being hydrolytically stable or bioavailable only at a pH below 8.
- A study comprising more than a year as performed for copper-based fungicides can be handled more easily indoors, provided that system size is sufficient and seasons can be simulated by light and temperature control.
- For focused testing under realistic worst case conditions of species known as particularly sensitive, the IME systems were shown to be the best option. Testing physical effects of paraffin oil layers at the water surface on insects breathing at the water surface, like heteropterans or culicidae, need sufficient size and water-sediment interactions to comprise appropriate planktonic communities and simulate degradation processes, respectively.
- For acaricides that inhibit the electron transport chain, extended fish studies combining the most sensitive species in flow-through juvenile growth tests (mostly rainbow trout) with realistic worst case exposure (static indoor microcosm) were developed.
- Due to the necessarily unique conditions of complex outdoor mesocosm studies, uncertainty factors are applied even to NOECs of well performed studies. The IME indoor microcosm systems were used for focused extensions of tested outdoor scenarios. The overall NOEC and NOEAEC of outdoor studies were tested/verified with a wide range of water and sediment qualities concerning nutrients and particle sizes and associated communities including added species of particular concern. It is also possible to extend the test conditions to temperature and light regimens of different climates and seasons.

Continuous exposure studies. For the purpose of setting water quality objectives or supporting metal risk assessments based on SSDs according to REACH it was necessary to perform community level studies with continuous exposure of an oligotrophic system as close as possible to realistic worst case conditions for bioavailability in Europe. Copper sulphate was dosed three times weekly to maintain the target concentration of dissolved copper in the water phase for three months. Sediment and biota were measured for total copper concentrations, water samples were measured for different copper species and the copper complexation capacity. The NOECs and LOECs were in line with the exposure phase of the copper fungicide study despite the different solubility of the applied copper salts. The effect thresholds tended to be lower than in known outdoor lentic mesocosm studies, but were clearly higher compared to lotic studies performed in small streams. When related to the concentration of free copper ions as effective copper species, effect thresholds were comparable. In conclusion, species sensitivities of stream or pond organisms most probably are not the reason for the higher sensitivity of lotic studies. It seems to be caused by the lower copper complexation capacity of small, fast flowing waters compared to ponds with a planktonic community and a comparably high DOC concentrations due to algae exudates.

Comparison of effects at different trophic conditions. The endpoint sensitivities of the exposure phase of the two indoor copper microcosm studies (copper-based fungicide study with a mesotrophic sediment and copper sulphate with an oligotrophic sediment) were compared to investigate the influence of trophic conditions on the study outcome. High nutrient levels caused by sediments with a high nutrient status and / or by spring conditions enable high population dynamics. At these conditions, intensive population growth amplifies direct effects on survival, growth and reproduction. At the same time, indirect effects by a lack of predation or competition become dominant and overrule direct effects on less abundant species, which might cause problems to show recovery in less complex indoor systems. Low nutrient levels caused by sediments with a low nutrient status and / or summer conditions limit population dynamics; populations exploit their habitat capacities at a seasonal climax stage. Direct effects on survival, growth and reproduction are less amplified and might be additionally masked by a reduction of intraspecific competition. Indirect effects will be less pronounced. At the same time, direct effects might occur at lower concentrations, as limited nutrition and competitive stress reduce the compensation potential. However, for a detection of these effects, a high statistical power is necessary. Thus, low nutrient levels should better be investigated in managed indoor microcosm studies.

Realization and communication of risk. In a project sponsored by the German UBA 2000/2001, we investigated by an interrogation, in how far the users of crop protection products also realize the environmental risks of use and exhibit creativity of risk mitigation. 1600 representative farms all over Germany, grouped in arable cropping small (up to 20 ha), medium (20-50 ha) and big farms (> 50 ha) and wine and orchard cultivating farms with 20% each were contacted by a call center and asked for an interrogation by a questionnaire with closed and open questions. 300 conversations could be evaluated. Between 21 and 27% of the contacted farmers of all groups except the small arable farms participated. Farmers in Baden-Württemberg and Bayern responded clearly over-proportionally, the eastern and northern Bundesländer exhibited the lowest response rates. The group of median arable crop farmers represented the biggest part of the participants in the interrogation with the highest estimated age and least dealing with issues of crop protection. Big arable crop farmers presented themselves as comparably professional in dealing with crop protection, and had the least critical view on potential environmental harm. Orchard farmers suffer most from the high quality pressure by the consumers. They realized a specific necessity of crop protection, resulting in high personal economic as well as high potential environmental loads. The actual practice of application regulations was judged clearly worst compared to the other groups. The group of the wine cultivating farmers was most communicative and creative. The contentment with the actual situation was clearly better than in the group of the orchard farmers.

Half of the responding farmers stated a possible relationship between the actual practice of crop protection and adverse environmental effects, mostly due to changed species distributions and deteriorated diversity by the use of herbicides. In all responding groups, there was acceptance of the general necessity of application regulations (80% of the wine and orchard-cultivating farmers, 50% of the arable crop farmers). Two thirds of the responding farmers regarded the actual regulations at least as acceptable, with orchard farmers clearly deviated from this trend. All groups ranked actual regulations being „optimal“ below 10%. Generally, co-operation between authorities and farmers and the communication of regulations should be improved, the notification process should be simplified and accelerated, the result should be more practice-oriented and the decisions should be more reliable in the end. Alternatives to chemical crop protection, such as crop rotation, mechanic measures, biological pest control, promotion of beneficial organisms, propagation of new thoughts, were frequently mentioned. There was a clear request for more research and development of biological pest control products. Politics and markets were perceived as actual hindrance of environmental protection, but also as possible tools for improvements in the future. Distortions of competition within the EU should be reduced. If conditions were equaled in all member states, regulations could be even more strict than today. The consumer should accept higher costs of sustainable agricultural production.

Water quality evaluation

Concepts. The EC Water Framework Directive sets priorities for hazardous substances and accelerates the derivation of water quality objectives. Chemical monitoring programs investigate whether the quality objectives for these substances are met or whether and how frequent critical values are exceeded as basis for the implementation of management measures. At the same time, water quality is evaluated to guarantee the identification of hazards caused by other influence factors than by the measured concentrations of a very limited number of substances. Indicator approaches comprise highly specific molecular biomarkers or receptors assay as well as indicator species for oxygen availability pointing to organic pollution. They are more integrative than chemical analyses, but cannot detect stressors outside their indicative potential. Habitat ecology approaches make use of habitat specific community structures by comparing them with the desired structure of comparable habitats of reference sites of highest pristine status possible.

> *Comparative assessment of national and international approaches to the classification of river health*. In 1996, before the implementation of the water framework directive, we asked water quality authorities and institutions of forty countries for detailed information on executively applied classification concepts with focus on chemical water quality. The interrogation comprised 30 European countries (14 EU member states and 16 non-EU-member states) and ten countries outside Europe, regarded as important due to population size and high technical progress. 23 countries responded to this inquiry, the percentage of 56-60% was similar in all three groups. The transferred information was entered into 24 descriptions of different approaches and the consecutive comparison. Thorough assessment of the validity of the transferred and processed information (origin and character of the information, completeness, status of national implementation) resulted in 18 approaches with sufficient validity. The classification approaches were described in order to clarify and compare the aspects Reference/Leitbild; objective of protection; quality elements and parameters; status, level and derivation of water quality objectives; classification system and reporting of classified river health.

Monitoring effects of pesticides in the field: Community monitoring studies. For the identification of pesticide-specific effects in the field it is necessary to exclude effects of other influence factors correlated with agricultural production, which may be anthropogenic (e.g. nutrient load, adjacent vegetation, water management measures) or due to agricultural site selection (e.g. local topography and hydrology). Benthic macroinvertebrate communities are widely used as indicators for water quality as they integrate potential impacts over considerable time periods. In the UK and Australia, a river invertebrate prediction and classification system was established, using reference sites classified according to their benthic macroinvertebrate community to predict the community of a monitoring site with comparable habitat characteristics. For the purpose of the presented study, we amended this list of well-established natural influence variables on macroinvertebrate community structure with variables specific for pesticide exposure and other agricultural influence factors to analyze the causality of potential pesticide exposure for changes of the macro-invertebrate community structure in two study areas. The “Altes Land” near Hamburg is characterized by intensive orchard cultures within a close web of ditches, representing a worst case of exposure to aquatic communities. The Braunschweig region is characterized by intensively cultured arable land in a diverse landscape, both aspects being more representative of situations in central Europe. For each agricultural region, 40 sampling sites were selected at different ditches to fulfill the general requirements of lowest possible variability in habitat characteristics, and widest possible range of potential pesticide exposure attributes. The sites were sampled five times for macroinvertebrates, water quality and other habitat characteristics from autumn 1998 to spring 2000. Macroinvertebrates were collected from sediment substrates (Braunschweig) or macrophyte and sediment substrates (Altes Land) and taxonomically determined at the species level where possible. Sites were clustered by their community structure to identify dominant influence factors.

Due to the unique situation of the sampling area “Altes Land”, most of the habitat variables typically affecting freshwater macroinvertebrate communities could be neglected: Especially topographical variables are very similar for the different sites. In addition, velocity (nearly zero) and the type of sediment were the same for all sites. Land use (grassland or orchard) was found to be the dominant factor influencing macroinvertebrate communities and explained more variability in community structure than the potential for pesticide exposure. Reasons for the special situation of the grassland ditches might be a minor shading, which might increase light intensity, water temperature and, consequently, primary production. The likely lower input of leaves into the grassland ditches might be one explanation for the tendency towards lower nutrient concentrations measured there, next to a potentially lower input of fertilizers in the extensively used grassland as compared to the intensively cultured orchards. Macrophyte densities did not seem to be related to grassland or orchards along the banks, but the species composition was different. The different terrestrial habitat structure close to the ditch might also be important for adults of insect species with aquatic larval stages. Focusing on the orchard sites only, the estimated potential of exposure was identified as the dominant gradient driving the community structure by principal component and redundancy analysis. The PRCs demonstrated the relation between potential of exposure and community structure over the seasons and identified *Assellus* and *Baetis* as the taxa that contributed most to the observed changes in the community structure. Clear and permanent effects on the communities could be demonstrated in the ditches with a high potential for pesticide exposure, which were very close to the trees (≤ 1.5 m). Differences between the sites with medium exposure potential (distance to the trees of 3 – 5 m) and the sites with a low (or negligible) exposure potential were generally smaller. The higher abundances of Pulmonata in the ditches with a high exposure potential might be explained by the usually low sensitivity of mollusks towards insecticides and the decrease of competition of more sensitive macroinvertebrates affected by the known significant insecticide inputs. The fact that Bivalvia were only found in ditches with a low or medium exposure potential, but not at potentially highly exposed sites, may be explained by enhanced susceptibility of early life stages and/or a lower recovery potential. Also caddisflies were not present in high-exposure sites. Our results show that a distance of 5 m between ditches and orchard trees should prevent significant long-lasting effects on the macroinvertebrate community in the “Altes Land”.

The presented methodology was also applied to the generation and evaluation of monitoring data in the Braunschweig region. Due to a much more variable landscape and use situation, we were not able to find 40 sites meeting the general requirements in a close area. The finally sampled sites were 20 sites south of Braunschweig (loess region), 12 sites north of Braunschweig (heath region) and 8 sites west of Braunschweig (Weser mountain region). A TWINSPAN analysis of community structure at all sites produced a first subdivision between nearly all Weser mountain sites and all the other sites. The second separation level produced one cluster comprising 75% of the heath and only 10% of the loess sites, and another with 85% of the loess and 25% of the heath sites. A canonical correspondence analysis showed the relation to the environmental factors, consistently explaining preferences of species forming communities in the subregions. The R- and H-values, indicating the geographical longitude and latitude, just hint to the western position of the most different Weser Mountain and the special northern heath communities. Highly differentiating properties are alkalinity, mainly spanned by the nutrient-rich loess sites, and stream velocity, mainly spanned by the mountain sites. The sediment structure is also characteristic with stones and gravel being correlated with community differences in the mountain sites and sand dominating the heath sites. Mud, loam and loess tend to indicate the loess site communities. Besides these geographical factors, depth and width of streams and ditches are very important for the community structure. As community patterns can be explained best by geographical subregions, it was decided that communities of the different subregions were investigated separately for pesticide effects to avoid mixing regional effects from effects of the (correlated) potential for pesticide exposure. Due to a lack of differentiation of the low potential of pesticide exposure, the Weser mountain sites were excluded from further evaluation. Braunschweig North and South sites in principle were well selected with respect to the range of potential exposure. However, the high losses of heath sites by drying up (75%) reduced the analysis to the first year and enabled a proper analysis of the full data for the loess sites only. We did not find correlations between community structure an the potential for pesticide exposure by PCA, PRC, species numbers or abundances of individual species. However, the calculation of the potential for pesticide exposure was much more uncertain compared to the “Altes Land”, as much more influence factors (geomorphology, water flow, upstream use) had to be included and the data was less exact.

Monitoring effects of pesticides in the field: Literature review. When focusing on monitoring of pesticides in the field, the objectives are moving towards the refinement of a substance related risk assessment. In a review of the most important published monitoring studies in Germany aiming at the effects of pesticides on non-target organisms we tried to classify objectives, approaches and methodologies and evaluate the reported effects with respect to the influence on the pesticide notification process. In all we analyzed 41 studies and classified them according to the investigated compartment (water body, terrestrial off-crop structures, in-crop soil), according to the methodology (chemical monitoring: total or biovavailable concentrations; ecotoxicological monitoring: active biological monitoring, bioassays with samples, molecular or physiological biomarkers; ecological monitoring, e.g. observation of effects on populations or species), and according to the main objective (water or soil quality, influence of agriculture, crop specific risk assessment, active substance related risk assessment, exposure (mitigation) of non-target areas). Most of the analyzed projects investigated pesticide fate and/or effects in water bodies, chemical monitoring being the main methodology. Ecological monitoring was performed with highly different intensity regarding the number of sites and frequency of sampling. It was partly added by ecotoxicological monitoring (bioassays). As a general conclusion from this review, thorough ecological monitoring is recommended in regions with high exposure potential, such as areas with a high intensity of pesticide applications (e.g. orchard cultures), with high run-off potential, or small catchment areas with intensive agriculture. In areas with special regulation, i.e., the „Altes Land“, the obligatory chemical monitoring should be supplemented by an ecological monitoring of sufficient sites to enable the separation of pesticide effects from that of other influence factors. When interpreting monitoring data it is helpful to include results of notification studies, especially higher-tier studies. However, data from the notification procedure are related to one product. The toxic unit concept can only be regarded as a very simple model to assess mixture effects. For assessing effects of culture-related pesticide use patterns, community level studies on the effects of those patterns are a clear improvement.

Artificial stream system for investigating habitat-specific fate and community effects. Experimental test systems simulating natural stream ecosystems with characteristic communities dominated by insect larvae are closer to the reality of edge-of the field streams and ditches than lentic mesocosms and may be an option especially for the environmental risk assessment of pesticides with insecticidal and adsorptive properties. A realistic simulation of pesticide loading should be based on an experimental system having sufficient size and complexity for self-preservation of the inherent populations at least for the duration of the study. Sediment composition and structure should be natural and inhabited by naturally age-structured populations. Investigations of fate and effects of the test substance(s) should be possible in different distances from the loading area, at sites with different stream velocity and in different microhabitats. Additionally it may be helpful that the system allows investigation of water leaching through the riverbed to simulate bank filtration processes and potential groundwater contamination. A realistic fate assessment should be based on studies at relevant concentrations, which may he very low, especially when the pesticide is highly toxic. As information about metabolites formation are important for the risk assessment, the use of radiolabeled substances is recommended. In the early 1990s, an appropriate artificial stream system with a total length of about 35 m, a width of 0.6 m (plus two wider pools) and a depth of 0.35 m was constructed at the Fraunhofer Institute in Schmallenberg. An electrically driven double chain with steel paddles is responsible for the generation of water flow with a stream velocity between 0 and 0.3 m/s. At the bottom of the concrete basin serving as containment, two aerated stainless steel tanks of 5000 L and 2500 L volume receive the contaminated water at the end of the exposure phase and serve as reservoir for uncontaminated freshwater, respectively. The concrete basin is filled with a 1 m drainage gravel layer and agricultural soil up to the upper margin of the trough. The facility is covered by a glass roof, which can open und close automatically via rain and wind sensors. Water and sediment including the original invertebrate community from a natural brook located in a water protection area are used to fill the artificial stream and model riffle sections. At first, the conditions of maintaining a complex stream macroinvertebrate community were investigated. Whereas the artificial stream system as an unicate system could not serve as stream mesocosm in the regulatory assessment of ecotoxicological effects, it can be used for experimental fate simulations and as reference system to validate hypotheses on effects on stream organisms under natural conditions.

Experiments using an adsorptive and biodegradable test substance demonstrate the capability of the system to simulate the property-specific temporal and spatial exposure pattern of substances, dependent on stream dynamics, sediment properties, and distance to the site of contamination. This can be of high interest for the specification of micro-sites for substance-specific sediment monitoring as well as for research on interactions between substance properties and site- and habitat-specific organisms and communities. The sediment surface of the artificial stream system can be structured to direct the water flow and create as different conditions as feasible at the same time. By using different labels (^{13}C, ^{14}C), different substances and their metabolites can be investigated at the same time.

Substance-related water quality evaluation concept based on a habitat approach. The habitat is the smallest environmental unit with defined structural and micro-climatic conditions, for which a typical biological community can be characterized. In aquatic systems, it is mainly determined by flow regimen, sediment structure, light conditions, and by the inhabiting community itself, especially primary producers and biofilms. The communities of structurally comparable habitats may differ due to water quality (including nutrient concentration and temperature), as well as due to potential species colonization patterns that may differ in different river basins and single water bodies, depending on connectivity, accessibility and the proportion of uncontaminated parts of the river net or water body.

The observation of a water body is performed at representative locations (sites). A water body consists of several habitats. A representative site comprises all characteristic habitats. Thus, for measuring ecosystem health of water bodies, a site-specific investigation integrating all habitats is demanded by the WFD and general practice in quality surveys. However, a generic assessment of substance-specific impact should focus on the individual habitats, as the concentration of pollutants in the sediment and on surfaces is dependent not only on the physico-chemical properties of the pollutant, but also on the habitat-specific surface structure and flow regimen (speed, turbulence), driving the contact rate of dissolved or dispersed materials with the surface. At the same time, the species characteristic for a specific habitat have specific needs being correlated with surface structure, flow regimen, and finally with the water exchange probability at respiratory membranes. Thus, habitat specific properties in combination with the feeding type are responsible for the uptake rate of pollutants. The concept of habitat specific fate and effects should be translated to and specified by a combined model (first step: habitat-specific fate and bioavailability calculation; to be combined with a kinetic uptake and distribution model and a MoA-specific effect model). This may serve as generic tool for the environmental risk assessment of chemical pollutants for all relevant water bodies to differentiate quality objectives according to the risk potential.

General conclusions and outlook to further research needs

After years of separate exposure and effect considerations in ecotoxicology, there is need to focus on exposure relevant for causing effects. This especially applies to the uptake processes in organisms. Bioavailability is a main key for understanding effects. For soluble substances, water exchange rates at respiratory membranes seem to be an important factor for uptake and thus internal exposure of target structures. Comparative toxicology needs the relation of effects to body doses or has to include information about uptake and elimination rates. For hardly water soluble substances and particulate matter, including nanoparticles and metal agglomerates of a diameter < 0.1 µm, ingestion and thus biomagnification is of higher importance and may explain e.g. phenomena of metal accumulation.

The universe of physiological variability in different taxa has to be taken into account. At the same time, a sensible generalization is necessary to focus on principles of relevant effects. In case of concern, specific modes of action need specific attention when planning and evaluating studies for hazard assessment. Targeted approaches (e.g. endocrine disruption test an assessment strategies) have to be set in ecological contexts and based on appropriate population-relevant endpoints. Behavioral reactions and interactions such as flight reactions and mating/spawning behavior should be accounted for when effects on populations in the field are to be predicted or interpreted.

With increasing level of biological system organization, complexity rises with functional redundancies and thus the capacity increases to compensate effects. The compensation capacity ends with the deterioration of accessible resources for running functions. At higher levels of systemic organization, concentration-effect-relationship become steeper as long as common regulative properties of the integrated levels are affected, which is true at least for the population level. Basing chemical risk assessments on higher levels of biological organization enhances the probablilty of including the most sensitive physiological interactions. However, NOEC values increase due to masking of physiological or individual effects by compensation. A steeper concentration-effect relationship increases the probability of severe effects on (sub-)populations at toxicant concentrations close to the observed effect threshold level.

To evaluate ecotoxicological influences in the field, the whole ecology has to be accounted for to identify and confounding factors and differentiate ecotoxiclogical from other effects. The central function of an ecosystem is to provide biodiversity. This function can most validly be evaluated by investigating the biological structure in community level studies. In the future, observations including population genetics may enhance ecological realism by identifying effects on the compensation potential for future hazards.

In a limited world with a growing human population and the challenge of a sustainable development of technical progress and management of resources, substance evaluations based only on hazard assessments often are not sufficiently realistic to accept uses of the substance. To achieve a pragmatic as well as protective regulation, substance evaluation should be based on risk assessment, open for a weight of evidence approach, aiming at a maximum of consistence. Especially in situations where benefits of environmental concentrations are opposed to risks (e.g., for pesticides and specific biocide uses or consequences of the use of building materials), the risks have to be assessed thoroughly regarding the most sensitive species and life stages at a realistic exposure profile of the most relevant habitats. The uncertainty of species sensitivity distribution should be quantified and evidence generated that there is no effect below a threshold concentration. The assessment can be proven or supported by monitoring studies, which should be focused on the influence factor of interest and the potentially confounding factors. The studies should be thoroughly planned to ensure detection of peak concentrations and concentration profiles as well as potentially sensitive endpoints regarding physiological, ecological and statistical aspects. For this, the observation methodology and the selection of sampling sites including appropriate reference sites are crucial.

The following fields of further research are presented in more detail.

Bioaccumulation. Whereas depuration seems to be comparable, the uptake mechanisms at respiratory membranes leading to bioconcentration and via ingestion leading to biomagnification differ considerably. Water solubility and/or adsorption potential of the test substance are responsible for the uptake route. This is actually considered in the revision of the OECD TG 305. For consumer protection, fish metabolism studies with pesticides are new requirements and under development. Bioaccumulation and fish metabolism studies require high numbers of vertebrates. Alternative test methods e.g. primary hepatocyte studies in combination with physiologically based kinetic modelling need to be developed.

From physiological to environmental effects. UNIFISH – a universal experimental tool for effect identification in fish is being developed using microarray approaches and microscopic high-resolution image analysis in zebrafish embryo tests to identify biochemical modes of action and phenotypic consequences of chemical stress. The fields of application may be non-target analysis of unknown contaminants in environmental or food samples, when an appropriate reference effect data base in available, or targeted identification of candidates for active substances in a high-throughput screening with whole organisms not regarded as animal tests by legislation. If a correlation of transkriptomic effects in embryos and sublethal effects in appropriate chronic studies can be established, UNIFISH can also be used for screenings on adverse side effects in fish (environmental hazard assessment) or vertebrates (human hazard assessment).

My experimental work started with studies on population dynamics, ecotoxicological endpoints in chronic fish tests and the attempt to find relations between individual toxic effects and population consequences. Now we can think about a combination of emission models, habitat-specific exposure modelling, physiologically based kinetic models, molecular effect models, individual effect statistics, predictive modelling of population effects and foodweb models for a thorough environmental risk assessment. As all of the models and combinations need validation by specially designed experimental setups, interesting research issues are coming up.

Population genetics and landscape level risk assessment. The population is the integrative biological level where predictive and retrospective objectives and methodologies of ecotoxicology and nature conservation meet. It is also the level at which processes of (micro-)evolution become manifest. The main factor influencing population genetics is intraspecific competition. The consequences of directed selection by anthropogenic stressors and gene flow between impacted and not impacted populations have to be taken into account when assessing the risk of chemicals applied to the environment. This requires regional and even local approaches.

Georeferenced risk characterization and prediction. The development of geodata and computing capacities enables probabilistic georeferenced approaches to risk assessment. We developed a novel approach for pesticide registration which is very promising, however, which still needs basic research input and opportunities for implementation. Finally, a vision of Integrated Risk Identification and assessment of Substances (IRIS) is presented.

Ausführliche Zusammenfassung

Ökotoxikologie umfasst Forschung und Regulation mit widersprüchlichen Zielen, Themen und Ansätzen. Ihr übergeordnetes Ziel kann definiert werden als **Identifikation und / oder Bewertung ökologischer Auswirkungen von Stoffen mit dem Anspruch der Minimierung von Unsicherheiten bezüglich Ursachen und Risiken**.

Aquatische Ökotoxikologie. Biochemische Reaktionen als die Basis für alle Lebensleistungen finden in wässriger Lösung statt und sind charakterisiert durch polare und hydrophobe Wechselwirkungen. Austauschprozesse von toxikologisch wirksamen Stoffen zwischen Umgebungsmedium und Organsimen sind am effektivsten zwischen dem Umgebungsmedium Wasser und der mit lipophilen Kompartimenten durchsetzten wässrigen Lösung in Organismen. Folglich besitzen aquatische Organismen mit hohen Wasseraustauschraten an äußeren Membranen das höchste Gefährdungspotenzial, wenn toxikologisch empfindliche molekulare Strukturen vorhanden sind. Aquatische Ökotoxikologie untersucht und/oder bewertet Wirkungen von Stoffen auf aquatische Ökosysteme. Der Forschungsschwerpunkt kann je nach Ziel variieren und unterschiedlichen Ansprüchen unterworfen sein. Im regulatorischen Kontext liegt der Anspruch entweder in der schlüssigen Bewertung eines zu registrierenden Stoffes, z.B. für die Zulassung seiner Anwendung in der Nähe von Gewässern, oder in der Bewertung der Wasserqualität, unter Umständen in Abhängigkeit von Typ und Nutzung des Gewässers. Für beide regulatorischen Anforderungen müssen wissenschaftsbasierte Ziele eingehalten werden, wie die Sicherstellung der physiologischen und taxonomischen Repräsentativität der Testorganismen hinsichtlich des Expositionspfades und der toxikologischen Empfindlichkeit, die Untersuchung oder Berücksichtigung des Kausalzusammenhangs zwischen Verursachung und Manifestation von Wirkungen, sowie der Kompensation von Wirkungen auf den verschiedenen Ebenen der biologischen Organisation wie Biomarker, Zelle, Individuum, Population oder Lebensgemeinschaft. Schließlich besteht die technische Herausforderung in der Entwicklung, Verbesserung und Standardisierung von Untersuchungsmethoden mit dem Ziel die regulatorischen Anforderungen wissenschaftlich angemessen zu erfüllen.

Stoffbewertung

Regulatorische Anforderungen. Stoffspezifische Daten zur aquatischen Ökotoxizität machen einen wesentlichen Bestandteil der benötigten Information zur Registrierung und Kennzeichnung von Stoffen und Produkten aus und können weitere Regulationen auslösen. Der Datensatz sollte protektiv für alle aquatischen Systeme sein, die einem Risiko ausgesetzt sind (Vorsorgeprinzip), aber sollte nicht zu überprotektiv sein, wenn ökonomische Interessen von Produzenten und Verbrauchern beschnitten werden. Deshalb sollte die Stoffbewertung am realistischen *worst case* orientiert werden. Als Grundlage für die Bewertung ist es notwendig, das aquatische Gefährdungspotenzial eines Stoffes durch Bestimmung seiner intrinsischen aquatischen Toxizität festzustellen. Im zweiten Schritt kann das Risiko für aquatische Systeme bewertet werden, indem die intrinsische Toxiziät zur Abschätzung der Expositionskonzentration in Beziehung gesetzt wird, welche sich je nach regulatorischen Zusammenhang (Pflanzenschutzmittelzulassung, Chemikalienanmeldung, Biozidzulassung, Zulassung von Veterinär- oder Humanarzneimitteln) deutlich unterscheiden.

Wegen begrenzter Ressourcen an Zeit, Geld und Expertenwissen ist es nicht machbar, alle chemischen Stoffe gleich gründlich zu untersuchen. So existieren in allen gesetzlichen Vollzügen gestufte Test- und Bewertungsverfahren, die mit ähnlichen Tests auf der Grundstufe starten. Sie sollen aquatische Systeme möglichst umfassend repräsentieren, um ein Maximum an möglichen Wirkungen zu geringstmöglichen Kosten detektieren zu können. Da die Repräsentativität ökotoxikologischer Testorganismen immer begrenzt ist, muss ein Sicherheitsfaktor eingeführt werden, um die Unsicherheiten der Extrapolation vom Basisdatensatz auf sichere Konzentrationen für die gesamte aquatische Lebensgemeinschaft abzubilden. Diese Unsicherheiten betreffen möglicherweise empfindlichere Arten, Lebensstadien oder Lebensleistungen als die getesteten, aber auch Unsicherheiten über indirekte Wirkungen im Nahrungsnetz. Mit steigender Exposition können die Unsicherheiten kritisch werden. Sie werden reduziert durch Tests mit zusätzlichen Standardtestarten oder durch chronische Tests. Falls die Umweltrisikobewertung keinen Sicherheitsspielraum zwischen geschätzter Umweltkonzentration und vorsorgender Toxizitätsabschätzung mehr aufweist, können höherstufige Testverfahren eingesetzt werden, um die Unsicherheit der Risikobewertung weiter zu reduzieren. So wird die Berechtigung zu erhöht, aquatische Ökosysteme einer entsprechenden Konzentration auszusetzen, da schädliche Wirkungen mit hoher Wahrscheinlichkeit ausgeschlossen werden können. Aus diesem Grund müssen die spezifischen Gefährdungseigenschaften des Stoffs oder Produkts identifiziert und in Studien geprüft werden, die sowohl die ökologische Realität widerspiegeln, als auch Repräsentanten der empfindlichsten Arten und Lebensstadien enthalten, sowie *worst case*-Bedingungen bezüglich der Exposition simulieren sollen. Die Einrichtung komplexer Testsysteme mit sensitiven Endpunkten und hoher statistischer Power bzw. geringer Variabilität, sowie die Aufrechterhaltung geeigneter Versuchsbedingungen über die gesamte Versuchsdauer erfordern ökologisches Verständnis, Erfahrung und technische Expertise.

Versuchsansätze zur ausführlichen Risikobewertung. Die beschriebenen ökologischen Versuchsansätze in der aquatischen Ökotoxikologie umfassen EU-Projekte, höherstufige Studien für die Industrie, F & E - Projekte für Regulationsbehörden und von Fraunhofer finanzierte Eigenprojekte. Die wesentlichen Themenfelder sind 1) die repräsentative Untersuchung und Interpretation von Art-Empfindlichkeitsverteilungen der besonders variablen Invertebraten, 2) die Berücksichtigung empfindlicher Lebensstadien und Lebensleistungen bei der Entwicklung von Full life cycle tests mit Fischen und die Aufstellung einer Teststrategie für endokrin wirksame Substanzen, 3) die Entwicklung einer *indoor*-Mikrokosmosanlage für die repräsentative Untersuchung von Wirkungen auf die aquatische Lebensgemeinschaft bei besonderen Gefährdungspotenzialen oder Expositionsmustern.

1) Die vorgestellten Projekte zur **Art-Empfindlichkeitsverteilung von Invertebraten** zielen auf physiologische Eigenschaften aufgrund taxonomischer und ökologischer Zusammenhänge, wie beispielsweise die Anpassung an die Bedingungen der spezifischen Habitate. Ein Vergleich der Art-Empfindlichkeiten von Bewohnern des Grundwassers oder von Oberflächengewässern gegenüber ausgewählten Pflanzenschutzmitteln zeigte, dass Inhibitoren des Aufbaustoffwechsels oder neuronaler Aktivitäten auf Grundwasserorganismen weniger stark oder mit deutlicher Verzögerung wirken. Es ergaben sich keine Hinweise auf eine physiologisch bedingt größere Empfindlichkeit der Grundwasserorganismen. Folglich wird die Empfindlichkeitsverteilung letaler Toxizität von Grundwasserarten hinreichend sicher abgeschätzt, wenn taxonomisch und physiologisch ähnliche Arten aus Oberflächengewässern getestet werden. Allerdings muss deutlich darauf hingewiesen werden, dass die Populationsdynamik von Grundwasserarten im Falle einer auftretenden Wirkung nur eine sehr langsame und unzureichende Wiedererholung durch Reproduktion oder Wiederbesiedlung leisten kann.

   Für Oberflächengewässerorganismen scheinen im Falle ubiquitärer toxischer Wirkmechanismen wie Acetylcholinesterase-Hemmung (gezeigt am Beispiel Carbaryl) die Empfindlichkeiten verschiedener Arten wesentlich von Aufnahmevorgängen und der davon geprägten Konzentration am Wirkort abhängig zu sein. Dabei spielt die Wasseraustauschgeschwindigkeit an respiratorischen Membranen eine wesentliche Rolle, die wiederum stark mit dem Sauerstoffbedarf korreliert und habitatspezifisch ist. Die vorgestellten Beispiele sind ein Beleg dafür, dass die Einbeziehung ökologischer Informationen die Bewertungsunsicherheit substanziell verringern kann. Um die Anforderungen einer ausführlichen ökologischen Risikobewertung zu erfüllen, sollten Art-Empfindlichkeitsverteilungen (SSDs) in zweiten Schritten genutzt werden: a) Einschluss von Arten aller relevanten taxonomischen Gruppen zur Absicherung, dass potenziell empfindliche Gruppen erfasst werden. Bei der Auswahl der Testarten sollten alle Informationen über Expositionsmuster und Gefährdungspotenzial genutzt werden. Die Steigung der Art-Empfindlichkeitskurve kann gering und die Kurvenanpassung unbefriedigend, die HC5 sehr niedrig, unsicher und damit von geringem Wert sein. b) Zusammenfassende Gruppierung / Auswahl der Arten nach taxonomischen, physiologischen oder habitatökologischen Kriterien. Einige Gruppen können wegen mangelnder Empfindlichkeit oder fehlender Exposition von der weiteren Auswertung ausgeschlossen werden. Die Risikobewertung für die verbleibenden empfindlichen Gruppen kann entweder nach Gruppen getrennt erfolgen und weitere Ansätze wie realistische Expositionsbetrachtungen einbeziehen, oder eine SSD über die kombinierten Daten der empfindlichen Gruppen nutzen. Die Steigung der Art-Empfindlichkeitskurve wird steiler, die HC5 höher und die Unsicherheit der Abschätzung niedriger. Allerdings sollten die niedrigsten Toxizitätswerte nahe oder sogar unter der HC5 darauf untersucht werden, ob mit dem statistischen Endpunkt schädliche ökologische Auswirkungen zu erwarten sind. Dabei sollten repräsentative chronische Wirkungen einbezogen werden. Falls die SSD auf EC50-Werten anstelle von NOEC- or EC10-Werten beruht, sollte die Exposition des fraglichen Habitats und das Wiedererholungspotenzial berücksichtigt werden. Das wird die regulatorische Unsicherheit der Gesamtbewertung über Fragen der statistischen Unsicherheit hinaus reduzieren.

2) Die vorgestellten Projekte zu **empfindlichen Lebensstadien und Lebensleistungen von Fischen** zielen auf die Detektion endokriner Wirkungen, Populationsrelevanz und Besonderheiten in der Regulation verschiedener Stoffgruppen. Auf der Basis der Erfahrung von 24 Fish Full life cycle tests (FLCT), Zwei- oder Mehrgenerationenstudien (2GT, MT) in öffentlichen oder vertraulichen GLP-Projekten, von denen 17 Tests endokrine Disruptoren mit verschiedenen Wirkmechanismen betrafen, wurden das Wssen um mechanismenspezifische Reaktionen vervollständigt. Endpunktspezifische Qualitätskriterien wurden abgeleitet und Testvorschriften und eine Teststrategie entwickelt, um endokrine Wirkungen erfassen und in ihrem Gefährdungspotenzial und Risiko bewerten zu können.

Für die abschließende Testung verdächtiger Substanzen wurden populationsrelevante oder apikale Endpunkte den für spezielle Wirkmechanismen indikativen Endpunkten gegenübergestellt. Populationsrelevante Endpunkte sind Schlupf, stadienspezifische Überlebensraten, Wachstum, Zeit bis zur ersten Eiablage (= Dauer der Sexualentwicklung), Geschlechterverhältnis, Fekundität (Eizahl) und Fertlilität (Befruchtungsrate). Indikative Endpunkte, wie sekundäre Geschlechtsmerkmale, Histopathologie, physiologische Biomarker oder Gonado-somatischer Index, sind nicht notwendig für die Risikobewertung, können aber wichtig sein zum Verständnis und zur Interpretation apikaler Wirkungen und ihrer Zuordnung zu Wirkmechanismen im Rahmen von Regulationsprozessen, die sich nur auf Gefährdungsklassifikationen gründen.

Besonders (aber nicht ausschließlich) in der Pflanzenschutzmittelzulassung muss die Auswirkung von kurzzeitigen Belastungsspitzen beurteilt werden. Testvorschriften für Langzeittests mit Fischen wurden jedoch für die Messung des intrinsischen Gefährdungspotenzials entwickelt und schreiben konstante Belastung vor. Wir entwickelten ein FLCT-Design mit dem Zebrabärbling für *worst case*-Belastungsspitzen in einem statischen Wasser/Sediment-System mit künstlichem sandigem Sediment. Drei verschiedenen Lebensstadien (Embryonen, Juvenile und Adulte) werden gleichzeitig einer einzigen oder wiederholten Anwendung der Testsubstanz im System ausgesetzt. Mit diesem einzigartigen Studientyp untersuchen wir das Potenzial der Testsubstanz und ihrer in Wasser oder Sediment gebildeten Metaboliten, während und nach einer Belastungsspitze Wirkungen auf die frühen Lebensstadien, die Juvenil- und Sexualentwicklung, die Reproduktion und die frühen Lebensstadien der Folgegeneration auszuüben. Da die drei Fischgruppen pro Testgefäß das gleiche Lebensstadium zu unterschiedlichen Zeiten nach der Anwendung und damit bei unterschiedlichen Konzentrationen der Testsubstanz durchlaufen, ergeben sich viele Vergleichsmöglichkeiten von Wirkungen zwischen den und innerhalb der Testreplikate. Durch Bezug der Wirkungen auf die gemessene Initialkonzentration können die Ergebnisse mit Abschätzungen der initialen Umweltkonzentrationen verglichen werden. Durch Bezug der Wirkungen auf zeitlich gewichtete Durchschnittskonzentrationen während empfindlicher Lebensstadien können die Ergebnisse mit intrinsischen Toxizitätswerten verglichen oder als solche verwendet werden. Die Endpunkte können nach ihrem Potenzial zur Wiedererholung und folglich der nachhaltigen Bedeutung von Wirkungen unterschieden werden. So zeigte bei Zebrabärblingen die Geschlechtsumwandlung von Weibchen zu Männchen keine Wiedererholung. Wirkungen auf das juvenile Wachstum oder die männliche Entwicklung einschließlich der Befruchtungsfähigkeit erholte sich nach Ende der relevanten Belastung, sowie bei geeigneter Ernährung und Konkurrenz im Zeitraum von Wochen bis Monaten, was von der Stärke der primären Wirkung abhing. Fekundität erholte sich innerhalb von Tagen, da bei den durchgeführten Studien die primäre Wirkung die Balzaktivitäten betraf, die wieder einsetzten, sobald die Belastung unter einen Schwellenwert gesunken war.

Für die Untersuchung sexualendokriner Wirkungen wurden die sich entwickelnden intrinsischen Testvorschriften für *in vivo-screening*-Tests (FSTRA, 21d-FA) und Langzeittests (FSDT, FLCT, 2-GT and MT) verwendet, um Wirkungen von Östrogenen, Anti-Östrogenen, einem Androgen und einem Anti-Androgen, Aromatasehemmern und anderen Inhibitoren der Steroidsysnthese zu identifizieren und zu bewerten. Die Versuchsergebnisse und deren Diskussionen mit dem UBA führten zu regulatorischen Überlegungen und Festlegungen auf eine Teststrategie, die eine Grundstufe zur Identifizierung endokrinen Potenzials (strukturelle Informationen und *read-across*, *in vitro*-Screening-Daten, Information aus toxikologischen Studien und Wirkungsstudien an Wirkstoffen), eine Stufe 1 zur Identifikation endokriner Wirkung bei relevanten Konzentrationen im Wasser (*in vivo*-Screening, 21d-FA) und eine Stufe 2 (abschließender Test für populationsrelevante Gefährdungsabschätzung) für ausreichend hält. Die Beiträge des IME sind:

Grundstufe. Falls das Verhältnis von akuten zu chronischen Wirkungen (ACR) bei Fischen hoch ist und falls kein anderer Wirkmechanismus dafür verantwortlich gemacht werden kann, ist der Verdacht einer endokrinen Wirkung nicht auszuschließen. Erfahrungen aus wissenschaftlicher und regulatorischer Testung legen einen ACR>20 als Auslöser nahe.

- *Stufe 1.* Da in gestuften Test- und Bewertungsansätzen die Möglichkeit falsch negativer Ergebnisse minimiert werden sollte, ist die von EDSTAC vorgeschlagene Strategie, drei *in vivo*-Stufen zu durchlaufen (Fish screening assay, FSTRA/Partial life cycle test und 2-GT) ungeeignet. Entweder erfasst der Stufe 2-Test (FSTRA/Partial life cycle test) die empfindlichsten Zeitfenster der Exposition und der Manifestation von Wirkungen und kann den definitiven Test ersetzen, oder nicht. Im letzten Fall führt er zu einem falsch negativen Ergebnis und zu einem falschen Ausstieg aus der Teststrategie. Der FSDT (OECD 234) kann als Endstufe-Test verwendet werden, wenn sichergestellt ist, dass das empfindlichste Expositionsfenster und das Stadium der empfindlichsten Effektausprägung erfasst sind. Nach heutigem Kenntnisstand kann das für Aromatasehemmer und Androgen-Rezeptor-Agonisten angenommen werden. Der FSTRA oder Partial life cycle test, die mit der Reproduktion adulter Tiere beginnen und spätestens mit den frühen Lebensstadien der Folgegeneration enden, sind nicht als abschließende Tests geeignet, da für fast alle sexualendokrinen Mechanismen gezeigt werden konnte, dass eine Exposition während Phase der Sexualentwicklung die empfindlichsten Wirkungen hervorruft, auch wenn diese erst in späteren Lebensphasen (Reproduktion, filiale frühe Lebensstadien) manifest werden.

  Bezüglich des Vorhersagepotenzials der Biomarker, die in Zebrabärblingtests des IME eingesetzt wurden, ist der VTG-Anstieg eindeutig indikativ für Östrogen-Rezeptor-Agonisten, während die VTG-Abnahme in Weibchen vermännlichende Wirkungen durch Östrogenrezeptor-Antagonisten und Aromatasehemmer anzeigen kann (Voraussetzung: keine vollständige Geschlechtsumwandlung!). Der Anstieg des in Fischen aktiven männlichen Hormons 11-kT deutet auf eine kompensatorische Reaktion auf Androgenrezeptor-Antagonisten hin. Die gleichzeitige Abnahme beider Biomarker lässt eine generelle Steroidsynthesehemmung vermuten, entweder durch eine Hemmung der Testosteronsynthese oder durch systemische Toxizität. Die in *in vivo*-Screenings (21d-FA, OECD 230) beobachteten Empfindlichkeiten liegen innerhalb derselben Größenordnung wie populationsrelevante Wirkungen in abschließenden Tests. Die einzige Ausnahme bildet der Androgenrezeptor-Agonist, bei dem die Biomarker-Antwort bei weitem weniger empfindlich ausfiel als die Geschlechtsumwandlung bei Exposition während der Sexualentwicklung. So sollten bei Verdacht auf Androgenrezeptor-vermittelte vermännlichende Wirkungen Testfischarten mit leicht unterscheidbarem Geschlecht und gut beobachtbaren sekundären Geschlechtsmerkmalen, die innerhalb von drei Wochen Exposition ansprechen, verwendet werden, zumindest bis ein geeigneter und empfindlicher Biomarker zur Verfügung steht.

- *Stufe 2.* Wir schlagen als abschließenden Test einen 2-GT vor, der mit befruchteten Eiern startet und nach der Sexualentwicklung der Folgegeneration endet. Im Vergleich mit dem durch Japan und die USA entwickelten Medaka MT (MMT) ermöglicht das Design die Exposition der empfindlichsten Lebensphase während der Sexualentwicklung in beiden Generationen und damit das Maximum an Möglichkeiten des maternalen Transfers von Wirkungen. Gleichzeitig reduziert das vorgeschlagene Design die Anzahl der verwendeten Fische um 17 %, indem es die anfängliche Reproduktionsphase auslässt, die für große Variabilität sorgt und bereits in den Screening-Tests untersucht wurde. Dennoch entwickelt das IME einen MT mit dem Zebrabärbling entsprechend dem MMT.

3) Der Wert, die Möglichkeiten und Beschränkungen von Ansätzen zur **Testung von Lebensgemeinschaften** für regulatorische Risikobewertungen wurden während des CLASSIC-Workshops diskutiert, der 1999 vom Autor in Schmallenberg organisiert wurde, und während des AMPERE-Workshop in Leipzig 2007 weiter spezifiziert. Wesentliche und teilweise zuwiderlaufende Aspekte von Lebensgemeinschafts-Studien sind a) Realismus bezüglich Exposition und Lebensgemeinschaften der Agrarlandschaft, b) Einschluss von Organismen mit Repräsentativität für die Arten, die sich bei Grundstufe-Tests mit der Testsubstanz oder aus der Literatur als die empfindlichsten erwiesen haben, c) Extrapolationsmöglichkeit auf eine große Spannbreite natürlicher Bedingungen. Unsicherheitsfaktoren, die auf die NOEC oder NOEAEC der Studie angewandt werden, müssen berücksichtigen, in wie weit die Versuchsbedingungen den realistischen *worst case* bezüglich der Exposition (Verteilung im und Verschwinden aus dem System) und der Empfindlichkeit der Endpunkte (Zusammensetzung der Lebensgemeinschaft, Anwesenheit empfindlicher Arten, funktionale Endpunkte) abbilden.

   Wir entwickelten semirealistische Mikrokosmossysteme von 1 m³ Volumen mit Sonnenlichtsimulation und regelbaren Temperaturen zwischen 5 und 40 ± 1°C zur repräsentativen Untersuchung von Wirkungen auf aquatische Lebensgemeinschaften. Ziel war die Fokussierung auf besondere Fragestellungen und Expositionsmuster und die Reduktion von Bewertungsunsicherheiten bei der Extrapolation von Freiland-Mesokosmosstudien auf andere Umweltbedingungen.

   Studien zur Zulassung von Pflanzenschutzmitteln. Es werden Beispiele für Studien in den IME-Systemen mit bestimmten Fragestellungen für Herbizide, Insektizide/Akarizide und Fungizide vorgestellt und im Vergleich zu Freilandstudien in Mesokosmen diskutiert. Das kleinere Volumen und die Testung im geschlossenen Raum ist ein klarer Nachteil bezüglich der Biodiversität der zu prüfenden Lebensgemeinschaft. Dominanzmuster werden starker ausgeprägt. Abundanzen größerer Makroinvertebraten oder Wasserpflanzen sind für eine statistische Auswertung meist zu gering. Diese Teile der aquatischen Lebensgemeinschaft sollten besser in größeren Freiland-Mesokosmen untersucht werden. Demgegenüber ist die Handhabbarkeit und die Konstanz der Bedingungen in den IME-Mikokosmen eindeutig überlegen.

- Daraus ergeben sich Vorteile bei der Testung von Stoffen die nur bei pH-Werten unterhalb 8.0 hydrolytisch stabil oder bioverfügbar sind.
- Eine Studie mit einer Dauer von über einem Jahr wie für Kupferfungizide kann im Gebäude leichter gehandhabt werden als im Freiland, vorausgesetzt dass die Systemgröße für die Dauer ausreichend ist und Jahreszeiten über Licht und Temperatur simuliert werden können.
- Für Lebensgemeinschaftsstudien unter *realistic worst case*-Bedingungen mit dem Fokus auf Arten mit aufgrund ihrer Lebensweise besonders empfindlichen Physiologien haben sich die IME-Systeme als beste Option erwiesen: Die Testung physikalischer Wirkungen von Paraffinölfilmen an der Wasseroberfläche auf luftatmende Wasserinsekten, wie Wasserwanzen oder Stechmückenlarven, benötigt genügend Volumen und Wasser-Sediment Interaktionen für die Beherbergung geeigneter Planktongemeinschaften und die Simulation natürlicher Abbauprozesse.
- Für Akarizide, die die Elektronentransportkette für Atmungsvorgänge hemmen, sind verlängerte Fischtests/Wachstumstests im Durchfluss mit der Regenbogenforelle häufig die empfindlichsten Studien, aber unrealistisch bezüglich der Exposition. In den IME-Mikrokosmen kann die empfindlichste Fischart mit einer *realistic worst case* – Exposition kombiniert werden.
- Wegen der grundsätzlich einzigartigen Bedingungskombination komplexer Freiland-Mesokosmen werden auch NOECs gut durchgeführter Studien mit Unsicherheitsfaktoren belegt. Die IME-Mikrokosmen werden für die gezielte Erweiterung getesteter Freiland-Szenarien eingesetzt. Die NOEC und NOEAEC von Freiland-Studien wurden mit einer größeren Spannbreite von Wasser- und Sedimentqualitäten hinsichtlich Nährstoffgehalt und Partikelgrößen getestet und verifiziert. Dabei wurden die mit den Sedimenten assoziierten Lebensgemeinschaften und weitere Arten eingesetzt, die im Verdacht standen besonders empfindlich zu sein. Ebenso ist es möglich, die Testbedingungen auf Licht- und Temperaturregime anderer Jahreszeiten oder Klimazonen einzustellen.

Studien mit kontinuierlicher Exposition. Für das Ziel der Ableitung von Wasserqualitätszielen oder der Unterstützung der auf SSDs basierenden Risikobewertung von Metallen nach REACH war es notwendig, Lebensgemeinschaftsstudien mit kontinuierlicher Belastung unter oligotrophen Bedingungen durchzuführen, die hinsichtlich der Bioverfügbarkeit so nah wie möglich an den *realistic worst case*–Bedingungen für Europa waren. Kupfersulfat wurde dreimal wöchentlich dosiert, um die Zielkonzentration an gelöstem Kupfer in der Wasserphase über drei Monate aufrecht zu halten. Sediment und Biota wurden auf den Gesamtgehalt an Kupfer analysiert, Wasserproben auf verschiedene Kupferspezies und die Kupferkomplexierungskapazität. Trotz der unterschiedlichen Löslichkeit der applizierten Kupfersalze entsprachen die NOECs und LOECs denen der Applikationsphase der Kupferfungizidstudie. Die Wirkschwellen lagen tendenziell niedriger als in den aus der Literatur bekannten Freiland-Mesoksomosstudien in stehenden Gewässern, aber deutlich höher als in den Studien, die in kleinen Fließgewässern durchgeführt wurden. Bei Bezug der Wirkungen auf die Konzentration an freien Kupferionen verschwand der Unterschied. Daraus lässt sich ableiten, dass die geringeren Wirkkonzentrationen in Fließgewässern nicht auf höheren Empfindlichkeiten von Fließgewässerorganismen beruhen, sondern durch die geringere Kupfer-Komplexierungskapazität der kleinen, schnell fließenden Gewässer zurückzuführen ist, während Teiche mit einer Planktonlebensgemeinschaft einen höheren DOC-Gehalt aufgrund der Algenexudate aufweisen.

Vergleich der Wirkungen unter verschiedenen Nährstoffbedingungen. Die Empfindlichkeit verschiedener Endpunkte der Belastungsphase der zwei IME-Kupfer-Mikrokosmosstudien (Kupferfungizidstudie mit mesotrophem Sediment und Kupfersulfatstudie mit oligotrophem Sediment) wurde verglichen, um den Einfluss der Närstoffbedingungen auf die Studienergebnisse zu untersuchen. Hohe Nährstoffgehalte in der Wasserphase, die durch Sediment mit einem hohen Nährstoffstatus und / oder Frühjahrsbedingungen hervorgerufen werden, ermöglichen eine ausgeprägte Populationsdynamik. Unter diesen Bedingungen verstärkt ein starkes Populationswachstum direkte Wirkungen auf Überleben und Wachstum/Reproduktion im Vergleich zur Kontrolle. Gleichzeitig werden indirekte Wirkungen durch das Fehlen von Fraßfeinden oder Konkurrenz dominant und überdecken direkte Effekte auf weniger abundante Organismen. Das kann in weniger komplexen Mikrokosmen zu Problemen mit der Demonstration von Wiedererholung führen. Niedrige Nährstoffkonzentrationen in der Wasserphase aufgrund nährstoffärmerer Sedimente und / oder Sommerbedingungen begrenzen die Populationsdynamik; Populationen pendeln sich an der Habitatkapazität ein und erreichen ein saisonbedingtes Klimaxstadium..Direkte Wirkungen auf Überleben und Wachstum/Reproduktion werden weniger verstärkt und können zusätzlich maskiert werden, da die innerartliche Konkurrenz abnimmt. Indirekte Wirkungen treten weniger stark hervor. Gleichzeitig können direkte Wirkungen bei niedrigeren Konzentrationen auftreten, weil limitiertes Futter und kompetitiver Stress das Kompensationspotenzial verringern. Allerdings ist für eine Detektion derartiger Wirkungen eine hohe statistische Trennschärfe notwendig. Deshalb sollten Studien mit niedrigen Nährstoffgehalten besser in regulierten und intensiv beobachteten *indoor*-Mikrokosmen stattfinden.

Risikowahrnehmung und -kommunikation. Im Auftrag des Umwelbundesamtes führten wir 2000/2001 eine Befragung durch um zu erfahren, in wie weit Nutzer von Pflanzenschutzmitteln auch Umweltrisiken dieser Nutzung wahrnehmen und Kreativität bei der Risikominimierung entwickeln. 1600 repräsentative Betriebe in ganz Deutschland, gruppiert nach kleinen (bis 20 ha), mittleren (20-50 ha) und großen (> 50 ha) Feldbaubetrieben, sowie Wein- und Obstbaubetriebe zu geplant je 20 % wurden über ein Callcenter kontaktiert und um Mitarbeit gebeten, einen Fragebogen mit geschlossenen und offenen Fragen zu beantworten. 300 Gespräche konnten ausgewertet werden. Zwischen 21 und 27 % der kontaktierten Landwirte aller Gruppen mit Ausnahme der Kleinbetriebe nahmen teil. Landwirte in Baden-Württemberg und Bayern waren deutlich überrepräsentiert, die östlichen und nördlichen Bundesländer wiesen die geringste Antwortrate auf. Die Gruppe der mittleren Feldbaubetriebe stellte die größte Gruppe der Teilnehmer dar mit dem höchsten geschätzten Durchschnittsalter und dem unsichersten Umgang mit Pflanzenschutzmitteln. Landwirte großer Feldbaubetriebe präsentierten sich als vergleichsweise professionell im Umgang mit Pflanzenschutz und hatten die unkritischste Wahrnehmung bezüglich potentieller Umweltschäden. Obstbauen leiden am meisten unter dem hohen Qualitätsdruck durch die Verbraucher. Sie sahen eine besondere Notwendigkeit zu Pflanzenschutzmaßnahmen, die sich in hoher persönlicher wirtschaftlicher, aber auch hoher potenzieller Umweltbelastung niederschlägt. Die aktuelle Praxis der Anwendungsbestimmungen wurde von den Obstbauern eindeutig am schlechtesten beurteilt. Die Weinbauern zeigten sich am kommunikativsten und kreativsten. Die Zufriedenheit mit der aktuellen Situation war deutlich besser als bei den Obstbauern.

Die Hälfte der antwortenden Landwirte bestätigte einen möglichen Zusammenhang zwischen der aktuellen Praxis des Pflanzenschutzes und schädlichen Auswirkungen auf die Umwelt, meist aufgrund veränderter Artzusammensetzungen und verarmter Biodiversität durch den Einsatz von Herbiziden. Von allen Gruppen wurde die generelle Notwendigkeit von Anwendungsbestimmungen akzeptiert (80 % der Wein-und Obstabauern, 50 % der Feldbauern). Die meisten aller antwortenden Landwirte bis auf die Obstbauern sahen die aktuellen Regulationen zumindest als akzeptabel an. Alle Gruppen bewerteten jedoch zu weniger als 10 % die Bestimmungen als optimal. Allgemein sollte die Kooperation zwischen Regulationsbehörde und Landwirten verbessert und der Zulassungsprozess für Pflanzenschutzmittel vereinfacht und beschleunigt werden. Die Anwendungsbestimmungen sollten sich starker an der Praxis orientieren und die Entscheidungen verlässlicher werden. Alternativen zum chemischen Pflanzenschutz, wie Fruchtwechsel, mechanische Maßnahmen, biologische Schädlingsbekämpfung, Förderung von Nützlingen und die Einbringung ganz neuer Ideen, waren vielfältig genannt. Mehr Forschung und Entwicklung im Bereich der biologischen Schädlingsbekämpfung wurde stark nachgefragt. Politik und Märkte wurden als wesentliche Hindernisse für den Umweltschutz wahrgenommen, aber auch als mögliche Werkzeuge für zukünftige Verbesserungen. Wettbewerbsverzerrungen innerhalb der EU sollten verringert werden; bei gleichen Bedingungen für alle Mitgliedstaaten könnten Regulationen sogar strenger sein als heute. Höhere Kosten nachhaltiger Landwirtschaft sollte der Verbraucher bereit sein zu tragen.

Bewertung der Wasserqualität

Konzepte. Die EU-Wasserrahmenrichtlinie setzt Prioritäten für gefährliche Stoffe und vereinheitlicht die Ableitung von Qualitätszielen. Chemische Monitoringprogramme untersuchen, ob die Qualitätsziele für diese Substanzen eingehalten, oder ob und wie häufig kritische Werte überschritten werden. Auf der Basis dieser Daten können Maßnahmen zum Flussgebietsmanagement ergriffen werden. Gleichzeitig wird die Wasserqualität mit dem Ziel bewertet, Gefahren durch andere Einflussfaktoren als die gemessenen Konzentrationen einer sehr beschränkten Anzahl von Substanzen festzustellen. Indikatormethoden reichen von hochspezifischen molekularen Biomarkern oder Rezeptortests bis hin zur Erfassung der relativen Abundanz von Indikatorarten für den Sauerstoffgehalt, welche den Grad organischer Belastungen anzeigen. Diese Messungen sind zeitlich und stofflich integrativer als chemische Messungen, sprechen aber auch nicht auf Stressoren außerhalb ihres Indikationspotenzials an. Biozönotische Ansätze nutzen die habitatspezifische Struktur der Lebensgemeinschaft, indem sie über einen Vergleich mit der Struktur entsprechender Habitate in möglichst unbeeinträchtigten Referenzgewässern eine Abweichung detektieren, die dann einer Ursache zugeordnet werden muss.

Nationale und internationale Ansätze zur Klassifizierung der Fließgewässerqualität. 1996, vor der Implementierung der Wasserrahmenrichtlinie, baten wir für die Wasserqualität zuständige Behörden in 40 Staaten um die Zusendung detaillierter Informationen über im Vollzug befindliche Klassifizierungskonzepte von Fließgewässern unter besonderer Berücksichtigung der chemischen Wasserqualität. Die Befragung umfasste 30 europäische Staaten (14 EU-Mitgliedstaaten und 16 damalige Nicht-Mitgliedstaaten) und 10 Staaten außerhalb Europas, die wir aufgrund ihrer Bevölkerungsgröße oder ihres technischen Fortschritts als bedeutend ansahen. Mit insgesamt 23 Staaten antworteten 56-60 % aus jeder Gruppe. Eine gründliche Untersuchung der Validität nach Ursprung und Art der Information, Vollständigkeit und nationaler Implementation ergab 18 Ansätze mit ausreichender Validität. Die Klassifizierungsansätze wurden beschrieben und auf folgende Aspekte vergleichend untersucht: Referenz/Leitbild, Schutzziele, Qualitätselemente und –kenngrößen, Status, Höhe und Ableitung von Wasserqualitätszielen, Klassifizierungssystem und Darstellung der klassifizierten Gewässergüte.

Erfassung der Auswirkung von Pflanzenschutzmitteln im Freiland: Biozönotisches Monitoring. Zur Identifizierung pflanzenschutzmittelspezifischer Wirkungen im Freiland ist es notwendig, andere Einflussfaktoren, die mit der landwirtschaftlichen Nutzung korreliert sind, auszuschließen, seien sie anthropogenen Ursprungs (z.B. Nährstoffbelastung, gewässerbegleitende Vegetation, Gewässerbewirtschaftung) oder verbunden mit der Standortselektion (lokale Topographie und Hydrologie). Lebensgemeinschaften benthischer Makroinvertebraten werden als Indikatoren für die Wasserqualität genutzt, da sie potenzielle Beeinträchtigungen zeitlich und stofflich integrieren. In Großbritannien und Australien wurde ein Vorhersage- und Klassifizierungssystem für Fließgewässer auf Invertebratenbasis etabliert. Es nutzt Referenzstandorte, welche nach charakteristischen Lebensgemeinschaften klassifiziert sind, um anhand der entsprechenden Habitatparameter einer Untersuchungsstelle die Lebensgemeinschaft vorauszusagen. Für die präsentierte Studie ergänzten wir die Liste bekannter natürlicher Einflussgrößen um Kenngrößen, die spezifisch für die Exposition gegenüber Pflanzenschutzmitteln und andere landwirtschaftliche Einflussfaktoren sind, um die Kausalität potenzieller Pestizid-Exposition für Änderungen in der Struktur der Makroinvertebraten-Lebensgemeinschaften in zwei Gebieten zu untersuchen. Das “Alte Land” bei Hamburg ist durch intensiven Obstbau in einem dichten Netz von Wassergräben charakterisiert und stellt damit eine *worst case*-Situation für die Belastung aquatischer Lebensgemeinschaften durch Pflanzenschutzmittel dar. Die Region um Braunschweig ist durch intensiven Feldbau in einer vielfältigen Landschaft charakterisiert, und repräsentiert damit eher die landwirtschaftliche Situation in Mitteleuropa. In jeder landwirtschaftlichen Region wurden 40 Probenahmestandorte an unterschiedlichen Gewässern ausgewählt, die die allgemeinen Anforderungen der geringstmöglichen Variabilität in Habitatkenngrößen, aber der größtmöglichen Spannbreite potenzieller Exposition gegenüber Pflanzenschutzmitteln erfüllen mussten. Die Standorte wurden zwischen Herbst 1998 und Frühjahr 2000 fünfmal beprobt, um die Makroinvertebraten, Wasserqualität und andere Habitatkenngrößen zu bestimmen. Makroinvertebraten wurden aus dem Sediment (Braunschweig) oder aus Makrophyten und Sediment (Altes Land) gewonnen und taxonomisch wenn möglich bis auf die Art bestimmt. Die Standorte wurden anhand der Struktur der Lebensgemeinschaft geclustert, um dominante Einflussfaktoren zu identifizieren.

Wegen der einzigartigen Situation des “Alten Landes” konnten die meisten Habitatkenngrößen, die Makroinvertebraten-Lebensgemeinschaften beeinflussen, vernachlässigt werden. Insbesondere die topographischen Kenngrößen hatten an den verschiedenen Standorten fast identische Werte. Dazu waren die Strömungsgeschwindigkeit (nahe 0) und der Sedimenttyp an allen Standorten gleich. Die Landnutzung (Grünland oder Obstbäume) stellte sich als der bedeutendste Einflussfaktor für die Makroinvertebraten heraus; die Kenngröße erklärte mehr Variabilität in der Struktur der Lebensgemeinschaft als das Expositionspotenzial gegenüber Pflanzenschutzmitteln. Gründe für die spezifische Situation der Gräben im Grünland liegen sicherlich in der geringeren Beschattung, die den Lichteinfall und die Wassertemperatur und damit die Primärproduktion erhöht. Der geringere Eintrag von Blättern mag eine Erklärung für die Tendenz zu niedrigeren Nährstoffgehalten in den Grünland-Gräben des “Alten Landes” sein, neben dem höchstwahrscheinlich geringeren Eintrag von Düngern in die extensiv genutzten Wiesen im Vergleich zu den intensiven Obstkulturen. Die Dichte der Wasserpflanzen in den Gräben schien nicht von der begleitenden Nutzung abhängig zu sein, allerdings war die Artenzusammensetzung verschieden. Die unterschiedliche terrestrische Habitatstruktur (Grünland oder Obst) mag für adulte Insekten mit aquatischen Lebensstadien von Bedeutung sein. Bei ausschließlicher Fokussierung der weiteren Auswertung auf die Obstbaumkulturen erwies sich in Hauptkomponenten- und Redundanzanalysen, dass das abgeschätzte Expositionspotenzial der dominante Gradient ist, der die Struktur der Lebensgemeinschaft bestimmt. PRCs zeigten die Beziehung zwischen Expositionspotenzial und Struktur der Lebensgemeinschaft über die Jahreszeiten und identifizierten mit Wasserasseln und Eintagsfliegen (Baetidae).die Taxa, die zu den wesentlichen Änderungen das meiste beitrugen. Deutliche und anhaltende Veränderungen der Lebensgemeinschaft zeigten sich in den Gräben mit einem hohen Pestizid-Expositionspotenzial, die eine geringe Entfernung zur ersten Baumreihe ($\leq$ 1.5 m) aufwiesen. Die Unterschiede zwischen Standorten mit mittlerer (Entfernung zur ersten Baumreihe von 3 – 5 m) oder niedrigem bzw. zu vernachlässigendem Expositionspotenzial waren generell geringer und nicht mehr signifikant. Höhere Abundanzen von Lungenschnecken in Gräben mit hohem Expositionspotenzial können mit der üblicherweise geringeren Empfindlichkeit von Schnecken gegenüber Insektiziden und der reduzierten Konkurrenz durch empfindlichere Makroinvertebraten erklärt werden, da bedeutsame Insektizideinträge belegt sind. Die Tatsache, dass Muscheln nur in Gräben mit niedrigem oder mittleren Expositionspotenzial gefunden wurden, aber nicht an potenziell hoch belasteten Standorten, ist möglicherweise auf die erhöhte Empfindlichkeit früher Lebensstadien und/oder geringeres Wiedererholungspotenzial zurückzuführen. Auch Köcherfliegen fehlten nur in den Gräben mit hohem Expositionspotenzial vollständig. Die Ergebnisse zeigen, dass ein Abstand von 5 m zwischen erster Baumreihe und Grabenrand signifikante langanhaltende Wirkungen auf die Struktur der Maroinvertebraten-Lebensgemeinschaft im Alten Land verhindert.

Die vorgestellte Methodologie wurde auch zur Erhebung und Auswertung von Monitoringdaten aus der Region Braunschweig angewendet. Aufgrund der erheblich vielfältigeren Landschaft und Landnutzung waren wir nicht in der Lage 40 Standorte in einer eng umgrenzten Region zu finden, die die allgemeinen Anforderungen erfüllten. Die letztendlich beprobten Standorte umfassten 20 Gewässer südlich von Braunschweig (lLössregion), 12 Gewässer nördlich von Braunschweig (Heideregion) und 8 Gewässer westlich von Braunschweig (Weserbergland). Eine TWINSPAN-Analyse der Lebensgemeinschaft aller Standorte zeigte zunächst eine Trennung der Weserbergland-Standorte von den anderen Standorten. Auf der zweiten Trennungsebene bildete sich ein Cluster aus 75 % der Heidestandorte und nur 10 % der Lössstandorte, während das zweite Cluster 85 % der Löss- und 25 % der Heidestandorte enthielt. Eine kanonische Korrespondenzanalyse wies den konsistenten Zusammenhang zwischen Habitatkennwerten und Präferenzen von Arten nach, die die Lebensgemeinschaft in den drei Subregionen charakterisieren. Die R- und H-Werte indizieren geographische Länge und Breite und weisen auf die westliche Lage der abweichenden Weserbergland-Lebensgemeinschaften und der nördlichen Lage der speziellen Heidegemeinschaften hin. Stark trennende Eigenschaften sind Alkalinität, besonders deutlich an den nährstoffreichen Löss-Standorten, und Strömungsgeschwindigkeit, vor allem wichtig für die Berglandstandorte. Ebenfalls charakteristisch ist die Sedimentstruktur: Steine und Kies sind korreliert mit Unterschieden der Lebensgemeinschaft im Bergland, Sand dominiert die Heidestandorte, Schlamm, Lehm und Löß sind am ehesten mit den Lebensgemeinschaften der Lößregion korreliert. Neben diesen geographischen Faktoren sind vor allem Tiefe und Breite der Bäche und Gräben bedeutsam für die Struktur der Lebensgemeinschaft. Da die Variabilität der Lebensgemeinschaften am besten mit der Zugehörigkeit zu den verschiedenen Subregionen erklärt werden kann, wurde entschieden, die einzelnen Subregionen getrennt auf Einflüsse des Expositionspotenzials gegenüber Pflanzenschutzmitteln zu untersuchen, um eine Vermischung von regionalen Einflüssen und möglicherweise korreliertem Expositionspotenzial zu vermeiden. Da im Weserbergland grundsätzlich ein sehr niedriges und wenig differenzierbares Expositionspotenzial vorlag, wurden die dort gelegenen Standorte von der weiteren Auswertung ausgeschlossen. Die Standorte in Braunschweig Nord und Süd waren im Prinzip gut ausgewählt, um eine weite Spanne des Expositionspotenzials zu repräsentieren. Leider reduzierte der hohe Verlust von 75 % der Heidestandorte durch Austrocknung die Auswertung auf die ersten beiden Probennahmen. Daher war eine solide Auswertung eines vollständigen Datensatzes nur mit den Lössstandorten möglich. Wir konnten keine Korrelationen zwischen Struktur der Lebensgemeinschaft und Pestizid-Expositionspotenzial finden, weder mit PCA und PRC, noch bei Artenzahlen oder Abundanzen einzelner Spezies. Allerdings war die Abschätzung des Expositionspotenzials erheblich unsicherer als im “Alten Land”, weil deutlich mehr Einflussfaktoren (Geomorphologie, Wasserabfluss, landwirtschaftlich Nutzung flussauf) berücksichtigt werden mussten und deutlich weniger exakt bestimmt werden konnten.

Monitoring von Pflanzenschutzmittelauswirkungen im Freiland: Literaturstudie. Wird Umweltmonitoring auf Pflanzenschutzmittel fokussiert, bewegen sich die Ziele des Monitorings in Richtung der Bestätigung oder Verfeinerung einer stoffbezogenen Risikobewertung. In einer Literaturstudie der wichtigsten publizierten Monitoringstudien in Deutschland, die auf die Wirkung von Pflanzenschutzmitteln auf Nicht-Zielorganismen zielen, klassifizierten wir Ziele, Vorgehensweisen und Methoden und bewerteten berichtete Wirkungen hinsichtlich des Einflusses auf den Zulassungsprozess von Pflanzenschutzmitteln. Insgesamt wurden 41 Studien analysiert und klassifiziert nach untersuchtem Kompartiment (Gewässer, terrestrische Off-crop-Struktur, Ackerboden), der Methodologie (chemisches Monitoring: Gesamt- oder bioverfügbare Konzentrationen; ökotoxikologisches Monitoring: Aktives biologisches Monitoring, Bioassays mit Proben, molekulare oder physiologische Biomarker; ökologisches Monitoring: z.B. Erfassung von Wirkungen auf Populationen oder Artabundanzen), und der Ziele (Wasser-oder Bodenqualität, Einfluss landwirtschaftlicher Nutzung, anbauspezifische Risikobewertung, wirkstoffspezifische Risikobewertung, Exposition von Nichtzielflächen). Die meisten der analysierten Projekte untersuchten Pflanzenschutzmittelverbleib und/oder –wirkungen in Gewässern; chemisches Monitoring war die häufigste Methode. Ökologisches Monitoring wurde in sehr unterschiedlicher Intensität betrieben, was die Zahl der Standorte und die Häufigkeit der Probennahme angeht. Es wurde teilweise von ökotoxikologischem Monitoring (Bioassays) ergänzt. Als Schlussfolgerung dieser Literaturstudie wird gründliches ökologisches Monitoring in Regionen empfohlen, die ein hohes Expositionspotenzial aufweisen, wie Gebiete mit großer Intensität von Pflanzenschutzmittelanwendungen (z.B. Obstbau, Hopfenbau), Gebiete mit hohem Run-off-Potenzial, oder kleine Gewässereinzugsgebiete mit intensiver landwirtschaftlicher Nutzung. In Sondergebieten wie dem „Alten Land“ sollte das obligatorische chemische Monitoring durch ein ökologisches Monitoring ergänzt werden, welches ausreichend Standorte einschließt, um die Wirkungen von Pflanzenschutzmitteln von denen anderer Einflussfaktoren zu trennen. Bei der Interpretation von Monitoringdaten kann es hilfreich sein, Ergebnisse aus zulassungsrelevanten Studien einzubeziehen, insbesondere Studien zur ausführlichen Risikobewertung. Allerdings beziehen sich Studienergebnisse aus dem Zulassungsprozess in der Regel auf einzelne Produkte. Das Konzept der Toxic units kann nur als sehr einfaches Modell zur ersten Abschätzung von Mischungseffekten herangezogen werden. Zur Abschätzung der Wirkungen ganzer Anwendungsmuster verschiedener Pflanzenschutzmittel sind kulturbezogene Studien mit ganzen Lebensgemeinschaften in Mesokosmen, die diesen Anwendungsmustern ausgesetzt werden, eine deutliche Verbesserung.

Künstliches Fließgewässersystem zur habitatspezifischen Untersuchung von Stoffverbleib und Wirkungen auf die Lebensgemeinschaft. Versuchssysteme, die Fließgewässer-Ökosysteme mit charakteristischen Lebensgemeinschaften simulieren, die von Insektenlarven dominiert werden, sind der Wirklichkeit von Fließgewässern und Abflussgräben in der Agrarlandschaft erheblich näher als klassische Mesokosmen. Insbesondere bei der Umweltrisikobewertung von adsorptiven Insektiziden können experimentelle Fließgewässer eine Option sein. Eine realistische Simulation einer Pflanzenschutzmittelbelastung sollte in einem System mit ausreichender Größe und Komplexität zum Selbsterhalt der enthaltenen Populationen über den Versuchszeitraum erfolgen. Das Sediment sollte in seiner Zusammensetzung und Struktur natürlichem Sediment entsprechen und von Populationen mit natürlicher Altersstruktur bewohnt werden. Untersuchungen zu Verbleib und Wirkung der Mittel sollte in verschiedenen Abständen von der Eintragsstelle an Stellen mit unterschiedlicher Strömungsgeschwindigkeit und verschiedenen Mikrohabitaten möglich sein. Gleichzeitig wäre die Untersuchung von Versickerung durch das Sediment hilfreich, um Uferfiltrationsprozesse und potenzielle Grundwasserbelastungen zu simulieren. Eine realistische Abschätzung des Verbleibs kann nur bei der Verwendung realistischer Konzentrationen erfolgen, die sehr niedrig sein können, vor allem wenn das Mittel sehr toxisch ist. Da die Erfassung von Metaboliten wichtige Informationen für die Risikobewertung liefern kann, ist die Verwendung radioaktiv markierter Testsubstanzen empfehlenswert. In den frühen 1990er Jahren wurde eine entsprechende künstliche Fließgewässeranlage aus Edelstahl mit einer Gesamtlänge von etwa 35 m, einer Breite von 0.6 m (mit zwei erweiterten Buchten) und einer Tiefe von 0.35 m im radioaktiven Kontrollbereich des Fraunhofer-Instituts in Schmallenberg eingerichtet. Ein elektrischer Doppelkettenantrieb mit Paddelblechen erzeugt eine Strömung zwischen 0 und 0.3 m/s. Am Grund des umgebenden Betonbeckens können zwei belüftete Edelstahltanks von 5000 L und 2500 L Volumen kontaminiertes Wasser am Ende der Expositionsphase aufnehmen und/oder als Reservoir für unkontaminiertes Frischwasser dienen. Das Betonbecken ist mit 1 m Drainagekies gefüllt, auf den Ackerboden bis zum Rand der Gewässerrinne aufgetragen ist. Die Anlage ist mit einem Glasdach versehen, das in Abhängigkeit von Regen und Wind automatisch öffnen und schließen kann. Wasser und Sediment werden mit der originalen Makroinvertebraten-Lebensgemeinschaft einem Bach im Bereich eines Wasserschutzgebiets entnommen, um das künstliche Fließgewässer zu füllen und engere Bereiche mit erhöhter Strömungsgeschwindigkeit zu modellieren. Zunächst wurden die Bedingungen untersucht, unter denen eine komplexe Lebensgemeinschaft aufrechterhalten werden kann. Das System ist wegen fehlender Replikation nicht als Fließgewässer-Mesokosmossystem für regulatorische Studien zu ökotoxikologischen Wirkungen geeignet. Es kann aber für die Simulation des Verbleibs von Stoffen und als Referenzsystem zur Validierung von Hypothesen von Wirkungen auf Gewässerorganismen unter natürlichen Bedingungen verwendet werden.

Untersuchungen mit einer adsorptiven und metabolisierbaren Testsubstanz zeigten die Fähigkeit des Systems, je nach physikalisch-chemischen Eigenschaften räumlich und zeitlich unterschiedliche Expositionsmuster zu simulieren, die von der lokalen Strömungsdynamik, den Sedimenteigenschaften und der Entfernung vom Eintragsort abhängen. Das kann von großem Interesse bei der Festlegung von Bedingungen für die exakten Probenahmestandorte für stoffspezifisches Sedimentmonitoring sein, aber auch für die Forschung an Interaktionen zwischen Substanzeigenschaften, Habitatexposition und habitatspezifischen Populationen und Lebensgemeinschaften. Die Sedimentoberfläche des künstlichen Fließgewässers kann strukturiert werden, um den Wasserstrom zu leiten und so unterschiedliche Habitate wie möglich zu schaffen. Mit unterschiedlichen Markierungen (^{13}C, ^{14}C) können verschiedene Substanzen und ihre Metabolite gleichzeitig untersucht werden.

Konzept zur Ableitung und Überprüfung stoffbezogener Wasserqualitätsziele mittels eines Habitatansatzes. Das Habitat ist die kleinste Umwelteinheit mit definierten strukturellen und mikroklimatischen Bedingungen, für die eine typische Lebensgemeinschaft charakterisiert werden kann. In aquatischen Systemen wird ein Habitat im Wesentlichen durch die Strömungsbedingungen, die Sedimentstruktur, den Lichteinfall und die Lebensgemeinschaft selbst, insbesondere die Wasserpflanzen und Oberflächenfilme, bestimmt. Die Lebensgemeinschaften strukturell vergleichbarer Habitate können sich aufgrund von Unterschieden in der Wasserqualität (inklusive Nährstoffgehalt und Temperatur) und aufgrund der Möglichkeiten zur Kolonialsierung durch artspezifische Zuwanderung unterscheiden. Letztere kann sich in verschiedenen Wassereinzugsgebieten oder einzelnen Gewässern je nach Vernetzungsgrad, Durchgängigkeit und dem Anteil unbelasteter Teilgewässer unterschiedlich darstellen.

Gewässeruntersuchungen werden an repräsentativen Standorten durchgeführt. Da sich ein Gewässer aus mehreren Habitattypen zusammensetzt, enthält ein repräsentativer Standort alle charakteristischen Habitate. So verlangt die EU-Gewässerrahmenrichtlinie für die Messung der Wasserqualität standortspezifische Untersuchungen, die alle Habitate integriert, was auch gängige Praxis in Qualitätsüberwachungen ist. Allerdings sollte eine generische Untersuchung und Bewertung von Einzelsubstanzen habitatscharf durchgeführt werden, da die Konzentration eines Schadstoffs im Sediment und an Oberflächen nicht nur von dessen physikochemischen Eigenschaften, sondern auch von der habitatspezifischen Oberflächenstruktur und dem Strömungsprofil (Geschwindigkeit, Turbulenz) abhängt, welches die Kontaktwahrscheinlichkeit von gelösten oder dispergierten Stoffen mit der Oberfläche bestimmt. Gleichzeitig haben Arten, die für bestimmte Habitate charakteristisch sind, spezifische Anforderungen, die mit der Oberflächenstruktur, Strömungsdynamik und letztlich mit der Wasseraustauschgeschwindigkeit an respiratorischen Membranen korrelieren. Somit be-stimmen habitatspezifische Eigenschaften in Kombination mit der Ernährungsweise die Aufnahmerate von Schadstoffen. Das Konzept habitatspezifischer Exposition und Wirkung sollte in ein kombiniertes Modell übersetzt werden, welches als generisches Werkzeug bei der Umweltrisikobewertung chemischer Belastungen in allen relevanten Gewässertypen Anwendung finden und die Qualitätsziele nach Risikopotential differenzieren kann. Dieses Modell könnte eine habitatspezifische Verbleib- und Bioverfügbarkeitsberechnung mit einem kinetischen Aufnahme- und Verteilungmodell und einem wirktypspezifischen Effektmodell kombinieren.

Allgemeine Schlussfolgerungen und Ausblick auf zukünftigen Forschungsbedarf

Nach Jahren der getrennten Betrachtung von Exposition und Wirkung wächst der Bedarf der Fokussierung auf die für die Wirkung relevante Exposition. Dabei spielen vor allem Aufnahmeprozesse in Organismen eine Rolle. Bioverfügbarkeit ist ein zentraler Schlüssel zum Verständnis von Wirkungen. Für wasserlösliche Substanzen scheinen Wasseraustauschraten an respiratorischen Membranen ein wesentlicher Faktor bei der Aufnahme und folglich der internen Konzentration an Wirkorten zu sein. Vergleichende Toxikologie benötigt den Bezug von Wirkungen auf Körperdosen oder muss Aufnahme- und Eliminationsraten einbeziehen. Bei schwer wasserlöslichen oder unlöslichen Stoffen inklusive Nanopartikeln ist die Aufnahme über den Darm und folglich Biomagnifikation von größerer Bedeutung und könnte möglicherweise Phänomene der Metallakkumulation klären helfen.

Die gewaltige Breite physiologischer Variabilität in verschiedenen Taxa muss berücksichtigt werden. Gleichzeitig ist eine vernünftige Generalisierung notwendig, um auf die Prinzipien relevanter Wirkungen zu fokussieren. Bei Verdacht müssen spezielle Wirkmechanismen bei der Planung und Auswertung von Studien für die Gefährdungsabschätzung besonders berücksichtigt werden. Auf besondere Wirkungen zielende Test- und Bewertungsstrategien sollten im ökologischen Kontext gesehen werden und auf geeigneten populationsrelevanten Endpunkten basieren. Verhaltensreaktionen oder Wirkungen auf Verhaltensinteraktionen wie Fluchtreaktionen oder Balz- und Laichverhalten sollten mit einbezogen werden, wenn Wirkungen auf Populationen im Freiland vorhergesagt oder interpretiert werden sollen.

Mit jeder höheren Organisationsebene biologischer Systeme steigt die Komplexität von Steuergliedern und –prozessen, die funktionale Redundanz und damit die Fähigkeit Wirkungen zu kompensieren. Die Kompensationsfähigkeit ist ausgereizt, wenn verfügbare Ressourcen verarmen, die für die Aufrechterhaltung von Funktionen notwendig sind. Auf höheren Ebenen systemischer Organisation werden Konzentrations-Wirkungs-Beziehungen daher steiler, so lange für alle auf der Ebene integrierten Prozesse allgemeine Regulationseigenschaften betroffen sind, was zumindest bis zur Populationsebene gilt. Wenn die Umweltrisikobewertung von chemischen Stoffen auf den höhere Ebenen der biologischen Organisation basiert, erhöht sich zwar die Sicherheit, dass die empfindlichsten physiologischen Interaktionen mit erfasst sind. Für diese erhöhen sich jedoch die NOECs auf höheren Ebenen, da physiologische oder individuelle Wirkungen durch Kompensation maskiert werden können. Eine steilere Konzentrations-Wirkungs-Beziehung stellt eine größere Gefährdung von (Sub-)Populationen bei Stoffkonzentrationen an der bestimmbaren Effektschwelle dar.

Um ökotoxikologische Wirkungen festzustellen und zu bewerten, muss die gesamte Ökologie einbezogen werden, da andere Einflussfaktoren identifiziert und ökotoxikologische Wirkungen von diesen Einflüssen unterschieden werden sollten. Die zentrale Funktion eines Ökosystems ist es, Biodiversität zur Verfügung zu stellen. Die Beeinträchtigung dieser Funktion kann am zutreffendsten bewertet werden, indem die biologische Struktur auf der Ebene der Lebensgemeinschaft untersucht wird. In Zukunft können Untersuchungen der populationsgenetischen Vielfalt die Beeinträchtigung der Kompensationsfähigkeit und damit das zukünftige Gefährdungspotenzial für Populationen realistischer erfassen.

In einer begrenzten Welt mit einer wachsenden menschlichen Population und den Herausforderungen einer nachhaltigen Entwicklung von technischem Fortschritt und Ressourcenmanagement sind Stoffbewertungen allein auf Basis einer intrinsischen Gefährdungsabschätzung immer weniger realistisch, um eine sichere Nutzung des Stoffs zuzulassen. Eine ebenso protektive wie pragmatische Regulation von Stoffen sollte auf einer Risikoabschätzung beruhen, die zusätzliche Informationen gewichtet einbezieht, um eine möglichst hohe Konsistenz zu erreichen. Besonders in Fällen, wo der Nutzen, der mit Konzentrationen in der Umwelt verbunden ist, den Risiken gegenübergestellt wird (z.B. bei Pflanzenschutzmittel- und Biozidanwendungen oder bei der Verwendung von Baustoffen), müssen die Risiken gründlich, d.h. im Hinblick auf die empfindlichsten Arten und Lebensstadien bei realistischem Expositionsprofil der relevanten Habitate, bewertet werden. Die Unsicherheit einer Art-Empfindlichkeits-Verteilung sollte quantifiziert werden. Dabei sollte möglichst eine Wirkschwelle festgelegt werden, unterhalb der eine Wirkung unwahrscheinlich ist. Die Risikoabschätzung kann durch Monitoringstudien überprüft oder unterstützt werden, die auf den zu überprüfenden Einflussfaktor (z.B. Konzentration eines Wirkstoffs im Gewässer) fokussiert werden und alle potenziell störenden Einflussfaktoren mit erfassen. Die Studien sollten gründlich geplant werden, um die Erfassung von Konzentrationsspitzen und Konzentrationsprofilen ebenso sicher zu stellen wie die Erfassung empfindlicher Endpunkte in physiologischer, ökologischer und statistischer Hinsicht. Das beinhaltet die Wahl der Erfassungsmethodik und der Probenahmestandorte einschließlich geeigneter Referenzstandorte.

Die folgenden zukünftigen Forschungsfelder werden detaillierter vorgestellt:

Bioakkumulation. Während die Ausscheidung sich ähnlich zu verhalten scheint, unterscheiden sich die Aufnahmemechanismen an respiratorischen Membranen, die zur Biokonzentration führen, und die mittels Nahrungsaufnahme, die zu Biomagnigfikation führen, deutlich. Wasserlöslichkeit und/oder Adsorptionspotenzial der Testsubstanz bestimmen den Aufnahmeweg. Dies wird gegenwärtig bei der Revision der OECD TG 305 berücksichtigt. Im Hinblick auf den Verbraucherschutz wurde durch Fisch-Metabolismusstudien mit Pflanzenschutzmitteln eine neue Anforderung formuliert, für die nun ein Testverfahren entwickelt wird. Bioakkumulations- und Fisch-Metabolismusstudien erfordern eine große Zahl von Wirbeltieren. Alternative Versuchsmethoden, etwa Studien mit primären Hepatozyten in Kombination mit physiologisch basierter kinetischer Modellierung, werden vorangetrieben.

Von physiologischen zu ökologischen Wirkungen. UNIFISH – eine universal einsetzbare Versuchsmethode zur Feststellung von Stoffwirkungen in Fischen - wird zurzeit entwickelt. Dabei werden Microarrays und hochauflösende mikroskopische Bildauswertung in Fischembryotests mit dem Zebrabärbling eingesetzt, um biochemische Wirkmechanismen und phänotypische Auswirkungen von Chemikalienstress zu identifizieren. Die späteren Anwendungsfelder können non-target-Analysen unbekannter Kontaminanten in Umwelt- oder Lebensmittelproben sein, sofern geeignete Referenzdaten für Effekte zur Verfügung stehen. Weiterhin können gezielt Kandidatensubstanzen in der Wirkstoffforschung im Hochdurchsatzverfahren mit vollständigen Organismen gescreent werden, ohne dass es sich dabei rechtlich um Tierversuche handelt. Falls Korrelationen zwischen transkriptomischen Effekten in Embryos und subletalen Wirkungen in geeigneten chronischen Studien aufgestellt werden können, kann UNIFISH auch für Screenings auf schädliche Nebenwirkungen auf Fische (Umweltgefährdung) oder Wirbeltiere (Humangefährdung) verwendet werden.

Meine wissenschaftliche Arbeit begann mit Studien zur Populationsdynamik und zu toxikologischen Endpunkten in Fischtests, verbunden mit dem Versuch, Zusammenhänge zwischen individuellen toxischen Effekten und Auswirkungen auf Populationen herzustellen. Heute können wir über Kombinationen zwischen Emissionsmodellen, habitatspezifischer Expositionsmodellierung, physiologisch basierten kinetischen Modellen, molekularen Wirkmodellen, individueller Effektstatistik, vorhersagende Modellierung von Populationseffekten und Nahrungsnetzmodellen für eine gründliche Umweltrisikobewertung nachdenken. Da all diese Modelle und ihre Kombinationen einer Validierung durch spezielle Studien bedürfen, sind interessante Forschungsthemen absehbar.

Populationsgenetik und Risikobewertung auf Landschaftsebene. Die Population ist die integrative biologische Ebene, auf der sich vorausschauende und feststellende Ziele und Methoden der Ökotoxikologie treffen. Sie stellt gleichzeitig die Ebene dar, auf der sich Prozesse der (Mikro-)Evolution manifestieren. Der größte Einflussfaktor für die Populationsgenetik ist die innerartliche Konkurrenz. Die Auswirkungen einer gerichteten Selektion durch anthropogene Stressoren und von Genfluss zwischen betroffenen und nicht betroffenen Populationen müsste bei der Risikobewertung von in die Umwelt ausgebrachten Stoffen berücksichtigt werden. Dazu braucht es regionale und möglicherweise sogar lokale Herangehensweisen.

Georeferenzierte Risikocharakterisierung und –voraussage. Die Entwicklung von Geodaten und Rechnerkapazitäten ermöglicht probabilistische georeferenzierte Ansätze zur Risikobewertung von Stoffen. So war das IME an der Entwicklung eines neuartigen Ansatzes für die Pflanzenschutzmittelzulassung beteiligt. Dieser ist vielversprechend, benötigt aber noch Einspeisung von Forschungsergebnissen und Gelegenheiten zur Umsetzung. Zuletzt wird die Vision einer Integrierten Risiko-Identifikation und –bewertung von Stoffen (IRIS) entwickelt.

1 Introduction

1.1 Science ruled by economy and politics

Ecotoxicology is dealing with effects of substances on life at different integrative levels of complexity in an ecological context. Depending on the scope of research, the objective is diagnostic, i.e., a retrospective identification of effects by monitoring a system, or prognostic, i.e., by conducting tests with a defined exposure enabling a prediction of potential effects (hazard identification).

Effects can be described as deviation from the normal status. Consequently, the normal status has to be known and clearly defined. Since there are considerable gaps of knowledge about the normal status of life at different integrative levels of biological complexity, especially when looking at aspects different from those touching human health, the first objective is the identification of effects (retrospective) or hazards (prospective) (ref. 28). If an effect can be identified and clearly attributed to a substance, the next objective may be different due to the research intention. Science is interested in the mechanisms being responsible for the effect, whereas producing companies, appliers and regulatory boards focus on the assessment of specific risks compared to benefits, and the derivation of risk mitigation measures. These are economical and political issues, but they also strongly interact with science since they drive financial support for research activities and dominate the orientation of cost intensive projects of high scientific quality. Since ecotoxicology has to relate effects to whenever possible quantified exposure, chemical analyses of more or less complex matrices aiming at partly very low concentrations have to be conducted. Since ecotoxicology has to consider ecological contexts, biological diversity and geographical confounding factors have to be accounted for. Good ecotoxicological science therefore has to be interdisciplinary and highly integrative with respect to expert knowledge and used methods. At the same time, regulative boards have to rely on the results, requesting quality assurance. The recent development goes forward to even research projects along or according to Good Laboratory Practice (GLP). Thus, modern ecotoxicological research can hardly be parted from regulatory objectives and industrial funding and is most successfully performed by a multidisciplinary professional team. The dependency of ecotoxicology on political scopes becomes evident in the different methods, views and general objectives represented by different expert panels (Table 1), which are mainly driven by different legal requirements. The overall objective of ecotoxicology may be defined as

Identification and / or assessment of ecological effects of chemicals, minimizing uncertainties of cause and risk

Table 1: Different scopes of ecotoxicological research

	Objective	
Substance evaluation	vs.	Environmental quality evaluation
	legal requirements by	
Chemical legislation, TGD	vs.	Water framework directive
Pesticide legislation, 91/414		Soil protection legislation
	aiming at	
Notification, registration	vs.	Quality objective surveillance
	Methodology	
Ecotoxicological studies	vs.	Monitoring

1.2 Specific risks for aquatic organisms and ecosystems (ref. 28), 50))

The aqueous solution represents the environment of all biochemical reactions. It is characterized by polar interactions of organic and inorganic molecules with water, and hydrophobic interactions or distribution of less polar organic molecules to non-polar compartments, such as protein carriers, storage lipids or membrane lipids as structural enclosures of cells. The uptake of toxicologically active substances by organisms happens by passing membranes as interfaces between organism and environment, in higher organisms i.e., membranes at leaf surfaces, skin, lung, gills or the intestinal tract, either by diffusion or by selective membrane transport processes. The unspecific uptake is characterized by

- the similarity of the environmental media to the biochemical environment,
- the properties of the chemical,
- the uptake properties of the membrane,
- the concentration of the chemical in the environmental media.

Exchange processes between water as environmental medium and the aqueous solution in organisms are the most effective ones. As the biochemical environment is considerably more lipophilic than the aquatic environmental medium, chemicals that are partitioned to cells are characterized by a certain water solubility and a wide range of lipophilicity, which can be described by a logarithmic partition coefficient n-octanol : water between 1 and 7 with the optimum between 3 and 6. Beside that, steric parameters and/or molecule size may be limiting factors for the uptake, depending on the pore size of uptake membranes.

The concentration gradient from the environmental medium to the organismic biochemical environment is a driving factor for the uptake of chemicals. Physiological performances can clearly enhance uptake. A high ventilation rate of the environmental medium results in constantly high concentrations at the uptake membranes. This ventilation rate can either be achieved by active suction and pumping of the medium, or by active movements relative to the medium, or by holding a fixed position in a flowing medium. On the inner side of the uptake membrane, a fast and steady removal results in constantly low concentrations and a high gradient. This can be achieved by high flow-rates of body fluids.

Consequently, organisms with the highest uptake potential of chemical pollutants are in close contact to a medium with high concentration of pollutants. They are characterized by high metabolic performance and resulting high oxygen needs, which is satisfied by a high ventilation rate of the environmental medium, a big respiratory membrane area, and a high-performance body liquid cycle. At the same time, these organisms are characterized by a maximum of the potential to metabolize and depurate chemical pollutants, and thus may but do not necessarily exhibit the highest bioconcentration. This profile is met best by vertebrates.

As air pollution is of highest risk for human health, terrestrial ecosystems profit from environmental health standards derived to meet that protection aim. In contradiction to that, humans use water for drinking rather than for breathing. Concentration standards derived for maximal daily intake dose are based on a drinking water passage of about 5% of the body weight through the intestinal tract. Aquatic organisms, especially fish, need passages of breath water at the gills being several orders of magnitude higher. In consequence, aquatic communities may not be sufficiently protected by drinking water standards and need specific risk assessments.

Different types of aquatic ecosystems are not only characterized by different communities, but also by different patterns of exposure to pollutants and different ways of exposure mitigation. Water soluble and adsorptive substances, bound to particles, are transported through watercourses. Lotic waters suffer most from exposure to waste water ingredients, which may be continuously released. At the same time lotic waters profit from connectivity enabling recolonization by migration from unpolluted tributaries. Isolated water bodies are lacking recolonization by aquatic life stages and any purgatory flow of unpolluted water. If polluted, they are more vulnerable. Freshwater ecosystems as a whole are very variable concerning physical and chemical water quality. The exact risk assessment of pollutants interacting with water quality parameters, i.e., pH, temperature, hardness or dissolved organic matter, needs a regional or even local perspective. Marine ecosystem conditions vary at much larger scales. However, as they are the final sinks of persistent aqueous pollutants, and as marine communities are much more diverse by means of taxonomy and physiology, they need specific protection. In Europe, marine regulatory risk assessment is only rudimentary developed so far.

1.3 Main focuses of aquatic ecotoxicology objectives

Aquatic ecotoxicology should identify or asses effects of chemicals on aquatic ecosystems. The focus of research, test development and evaluation strategy is dependent on the objective of result application. In summary, the focus can be oriented at the following aspects, which represent different dimensions interlinked in different ways (Figure 1):

- The regulatory objective: Substance evaluation or water quality evaluation
    - If substance evaluation: Context of competent legislation, i.e., for general chemicals, pharmaceuticals or pesticides
    - If water quality evaluation: Type (and use[1]) of water body, i.e., groundwater, lentic or lotic surface waters, or marine water
- The objective of representativity and extrapolation of research results
    - Taxonomic or physiological representation
    - Representation of communities subjected to specific paths of exposure, i.e., by municipal waste water effluents, pesticide spray drift or leaching to the groundwater
- The objective of investigating causal relationships: Effects and effect compensation on different levels of biological organization, i.e. biomarkers, cells, individuals, populations or communities
- The technical objective of developing, improving and standardizing methodologies, i.e., sensitive non-standard species tests, species-sensitivity distributions (SSDs), life cycle tests, modeling, specific microcosm studies, monitoring approaches.

[1] Water quality standards are also derived for goods of protection different from aquatic communities, such as irrigation water or drinking water. These standards are based on the protection aim of human health (either directly or *via* food production) and thus are outside the focus of this thesis. However, they have to be mentioned, because they are often not clearly differentiated or mentioned representative or protective for aquatic communities, which is not scientifically justified (see chapter 1.2).

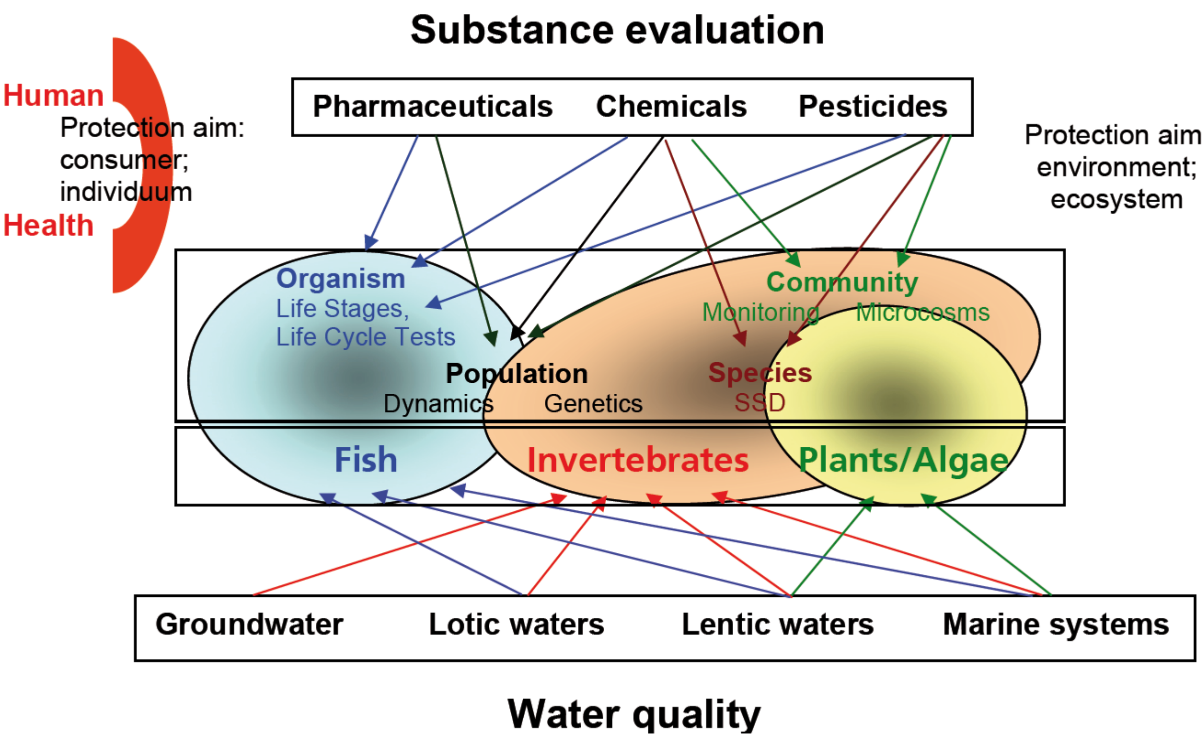


Figure 1: Overview of main focuses of aquatic ecotoxicology: Main objectives; legally differentiated substances and protection aims; biological integration levels, methods and organism groups; water types

1.4 Evidence of experimental results

When assessing ecological effects of chemicals, in most cases it is only possible to detect effects on endpoints focused at in predictive tests without attributing a precise causal relationship to a mode of action (ref. 28). Compared to the situation in classic toxicology, there is still more weight on the development of methods for effects detection and assessment than on the investigation of the processes contributing to ecological toxico-dynamics. The latter can only be validly approached by including advanced knowledge on systemic relationships and regulatory processes within and between the integrative levels of biological organization. However, even without that knowledge there is a practical need for risk assessments simply based on observed effects. The predictive power of these assessments strongly depend on the observation methods and the study conditions, for example

- the examined integrative level of biological organization,
- the test species,
- the endpoint criteria,
- the duration of exposure and observation.

Before designing a study, the objective and quality of the study have to be clearly defined: Optimized reproducibility and a high degree of standardization can hardly be in line with high ecological realism (Figure 2); a high predictive power for individual sensitive species or communities is different from a general functional representativity. Each objective requests an approach to specific aspects of toxicity, and consequently may need specific tests at suited integrative levels of biological organization. Thus, at the actual state of knowledge, it makes no sense to describe ecotoxicity as a generally understood generic process. Ecotoxicity can only be approached via its observability, i.e. the study types which to date are in use.

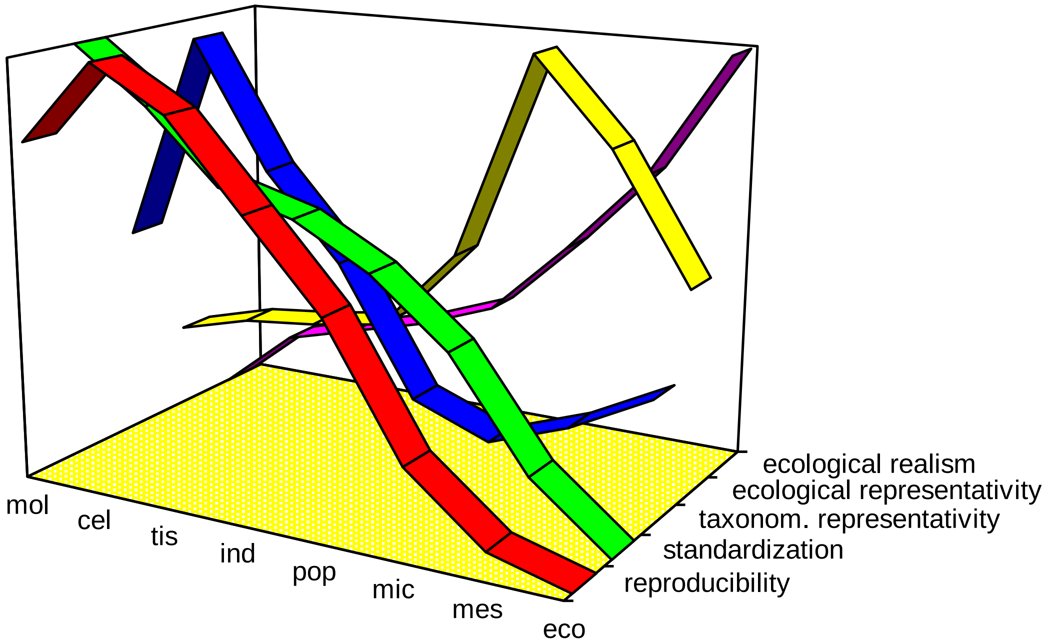


Figure 2: Evidence of experimental results at different levels of biological complexity with respect to different properties
mol = biochemical assays; cel = cell culture assays; tis = tissue related assays; ind = tests on individual organisms; pop = population studies; mic = microcosm studies, mes = mesocosm studies, eco = investigations in real ecosystems (ref. 28)

In this thesis, conceptual and experimental results are presented of performed research projects addressing ecological aspects of ecotoxicological risk assessment. Those projects include developments for substance-specific higher-tier aquatic risk assessment (substance evaluation, chapter 2) as well as for ecological monitoring approaches (water quality evaluation, chapter 3). Finally, conceptual outlooks for future research are given.

2 Substance evaluation

2.1 Introduction: Character of legal requirements and needs for ecological approaches

2.1.1 General principles and differences of competent regulations

Substance-specific aquatic ecotoxicity is a substantial part of information needed for registration and labeling of chemicals and products, triggering various further regulations, i.e., for setting security requirements for production plants and storage facilities, transport classification, restrictions in use, or prescriptions of waste treatment. The data set being used for the assessment of aquatic ecotoxicity as well as the assessment itself has to fulfill different requirements. It should be protective for all aquatic systems at risk (precautionary principle), but it should not be too overprotective in order not to overrule economic interests of producers and consumers. For these reasons, the assessment should be oriented at the realistic worst case. At the same time, it should be equal for all substances dealt with by the competent legislation and enable fair comparisons between different substances to minimize risk/benefit ratios of uses by choosing the best alternative. Thus, the data has to be of comparable quality and generated according to defined standards. As a basic step, it is necessary to assess the aquatic hazard potential of a chemical by determining its intrinsic aquatic toxicity. In a second step, the risk for aquatic systems may be assessed by comparing the intrinsic toxicity with estimations of exposure.

The main differences between the competent legal regulations are due to exposure. In pesticide regulation, exposure has to be regarded in principle as an unavoidable side effect of intended use in the field. The risk of a specific application of a specific product is to be minimized by mitigation measures, which focus on minimization of exposure. Only if the intrinsic aquatic toxicity is so high that adverse effects cannot be excluded by a maximum of accepted exposure mitigation measures, the specific application cannot be registered. Consequently, enormous efforts are spent in predicting realistic worst case exposure as precisely as possible for specific use patterns and climatic scenarios, taking into account the main exposure pathways spray drift, runoff and drainage. The primary water bodies at risk are edge-of-field drainage or irrigation ditches, small streams and ponds in the agricultural landscape.

In contradiction, industrial chemicals and chemical products as well as pharmaceuticals (and pesticides beside intended use) are not intended to be present in the environment. They mainly enter water bodies by presence in effluents or by insecurities or accidents of production, storage and transport. Thus, the primary water bodies at risk are canals, rivers and lakes with adjacent industry or trade facilities, shipping, and/or receiving effluents. However, when looking at worst case scenarios, also small and isolated water bodies may be affected by accidents on adjacent roads or crossing bridges. Effluents of sewage treatment plants receiving the same amounts of a specific pollutant may differ in concentration by orders of magnitude depending on the treatment concept and technology. For these reasons, a precise prediction of realistic worst-case exposure is neither possible, nor is it necessary. For industrial chemicals, the information on intended yearly produced tons is used as a rough exposure classification in the EC. With the implementation of REACH (Registration, Evaluation and Assessment of Chemicals), this is added by a classification of substances of very high concern (SVHC) due to very high persistence and bioaccumulation potential, or high toxicity in combination with persistence and bioaccumulation. If these properties or equivalent concerns such as endocrine disruption are identified, specific regulations are to be enacted, including authorization and restriction of uses. For human pharmaceuticals, producers have to indicate the aimed part of the market for specific active substances to enable exposure calculations.

2.1.2 Dealing with uncertainties: Tiered approaches

Due to the limited resources of time, money and expert knowledge, it is not useful to test all chemical substances in depth. Thus, in all legislative contexts tiered approaches have been implemented starting with similar tests at the lower tiers. These tests are chosen to ensure maximum representativity for detecting potential effects on aquatic systems at lowest costs.

The basic aquatic standard test set comprises tests with a unicellular green algae species, a cladoceran species (*Daphnia magna)*, and a fish species. These species represent the taxonomic groups (with related physiological properties) of plants (photosynthesis), arthropods as most important invertebrate group, and vertebrates. At the same time, these organisms have different positions in the food web of aquatic ecosystems, representing primary producers (phytoplankton), primary consumers (zooplankton) and secondary consumers, respectively.

The basic tests are short-term tests with test duration between 48h (acute *Daphnia magna* test) and 96h (acute fish test). The toxicological endpoints are mortality in fish (or immobility in daphnids) and cellular growth in algae, the statistical endpoint is the EC50 (effect concentration reducing abundance by 50% compared to the untreated control), as this is commonly the least variable and most reliable effect percentile due to the commonly normal distribution of effects. Following the precautionary principle, the most sensitive test result is used for hazard assessment.

When assessing the risk of a chemical for aquatic communities, the lowest effect concentration is compared with the appropriate predicted exposure concentration (PEC). As the representativity of the toxicity results is limited (ref. 38), an uncertainty factor has to be applied to address the uncertainties of extrapolation from the basic test set data to safe concentrations for aquatic communities, such as

- uncertainties about more sensitive species,
- uncertainties about potentially more sensitive life stages and performances (chronic effects),
- uncertainties about secondary effects.

Either the effect concentration is directly divided by the uncertainty factor to derive a PNEC (predicted no effect concentration), which must not be exceeded by the PEC, i.e., the PEC/PNEC ratio has to be <1. Or, as it is practice in the EC pesticide regulation, the toxicity is divided by the PEC. The resulting toxicity/exposure ratio (TER) should be higher than a trigger value representing the uncertainty of the assessment (SANCO 2002).

The general principle of tiered risk assessment approaches to is to relieve the evaluation process from all substances of minor toxicity (= hazard potential) and/or minor exposure potential. The EU chemicals regulation (REACH) classifies the exposure potential by production volume. At low production volumes (<10 t/a) and thus low exposure potential, only short-term toxicity studies on algae and invertebrates are required. However, low water solubility or high acute *Daphnia* toxicity may trigger the performance of a chronic study (*Daphnia* reproduction test, OECD TG 211). When exceeding a production volume of >10 t/a, fish studies become necessary. High hazard potential triggers a chronic fish study (e.g. a juvenile growth or early life stage test, OECD TG 215 or 210, respectively). More than 100 t/a require a chronic fish test anyway. The resulting data of the hazard assessment are used to assess the environmental risk by comparing the hazard potential with a PEC calculation.

The uncertainty factor for chemicals is dependent on the amount, diversity and quality of ecotoxicity data. It may range from 1000 (based on short-term EC50s or LC50s only) to 10 (based on NOECs in chronic tests with *Daphnia* and fish). If the PEC/PNEC ratio is <1, no refinement of the risk assessment is necessary and the evaluation is finished at the lowest tier. If a PEC/PNEC-ratio of 1 is surpassed (especially at high production volumes of >1000 t/a), or if special concerns arise (e.g. by a suspicion of endocrine disruption), the quality of the data set can be improved by testing more species or performing specific studies addressing the concerns (e.g. fish full life cycle studies for potentially endocrine substances). This helps to reduce uncertainty and thus increases the PNEC. Alternatively or additionally, refined exposure estimation or monitoring results may reduce the PEC.

Within the EC pesticide regulation (SANCO 2002), the TER-values triggering the need of a refined risk assessment are <100 and <10 for acute and chronic risk assessment, respectively. A reduction of uncertainty by additional information may reduce the TER-value acceptable for the registration of a specific use under specific use conditions. The acceptance is based on expert judgment and may be discussed in expert panels.

2.1.3 Ecological approaches to aquatic ecotoxicology

The main objective of higher tier risk assessment studies is to reduce uncertainty of the assessment and enhance the justification to expose aquatic ecosystems to predicted concentrations, as adverse effects can be excluded with high probability. For this purpose, the specific concerns of the chemical product have to be identified and addressed in the studies, which should include ecological realism as well as representation of sensitive species, life stages and worst case exposure conditions. Own study concepts were summarized in Schäfers and Klein, 1999 (ref. 12), matching well with the outcome of the HARAP workshop (Campbell et al. 1999). These studies should be as close as possible to standardized procedures, but include the special approaches needed to answer the open questions by the specific risk assessment. As these studies are tending to be unicates, they need much effort and money. They are only necessary and economically feasible for chemicals or products of high interest for the users and consumers. While triggered by economic interests, higher tier studies present challenges for ecological research and development:

- The preparation of complex test systems with sensitive endpoints and high statistical power/low variability and their maintenance for the time of investigations needs ecological understanding, continuity and technical expertise. It is a pre-requisite for reducing uncertainty.
- A defined (and quality assured) exposure to a stressor under controlled conditions enhances understanding of exposure situations of different aquatic habitats and related species. It also enhances understanding of effects by stressors on different levels of biological organization including consequences for the next levels, as well as mechanisms of effect compensation and recovery. These scientific aspects help to restrict the sensitive species to that at a realistic probability of exposure and to interpret causalities to generate consistent overviews of effect and recovery patterns.

In the following, the main projects of the last ten years representing ecological approaches to aquatic ecotoxicology are shortly described, including EC projects, higher tier studies for industry, R & D projects for authorities and own projects supported by Fraunhofer:

- the representative investigation of species sensitivities, the comparison of risk probabilities, the need to include information about evolution, physiology, and habitat ecology as well as evidence from mesocosm studies and monitoring data (chapter 2.2),
- the development of specific fish full life cycle test designs for different objectives focusing on the most sensitive life stages and performances, the extrapolation to fish populations, the investigation of biomarkers and the derivation of a fish test strategy for endocrine disrupting chemicals (chapter 2.3),
- the development of an indoor semi-realistic microcosm system for the representative investigation of community effects regarding specific concerns and exposure patterns, as well as for the performance of extended laboratory studies at realistic

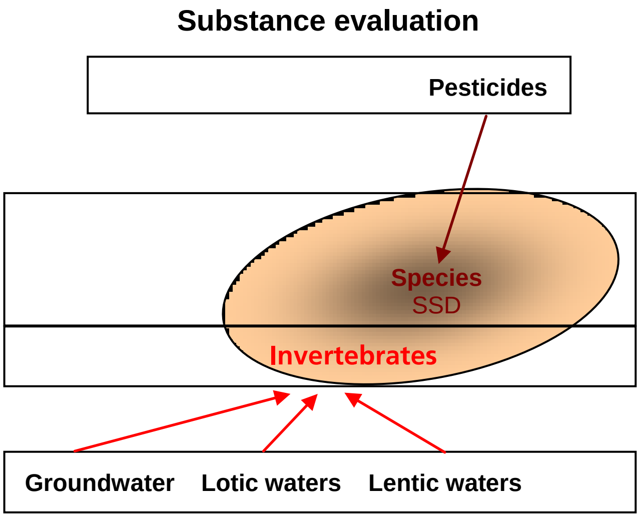
worst case exposure; influences of nutrient state and season; macrophyte studies (chapter 2.4),

- and finally coming back to the central trigger of precise ecological risk assessment, the consumer needs: an interrogation of pesticide users concerning the realization of ecological effects, risk mitigation and consumer pressure (chapter 2.5).

The last chapter does not describe mainly ecotoxicological research projects, but projects and proposals that needed experience in ecological risk assessment. It is regarded important to understand the social implications of decisions derived from ecotoxicological data and the necessity for risk communication to the end-users and consumers, not least to convince society and politicians to further invest in ecological approaches to ecotoxicology.

2.2 Sensitive species: Invertebrates

2.2.1 Introduction



Species sensitivity distributions (SSDs), based on effect probability quantiles, are increasingly accepted by regulatory authorities to reduce uncertainties of species extrapolations (EUFRAM 2006)[2].

From a scientific view it clearly preferable to quantify data variability and include uncertainty in the risk assessment by calculating hazard, exposure or risk probabilities rather than addressing it by scientifically not justified uncertainty factors, based on Arabic mathematical tradition (decimal system) or gut feelings.

Unfortunately, this approach is often used in a simplified technical manner without concept of substantial data quality. From a regulatory view, the development of SSD approaches should move the focus from the best probabilistic statistics to the appropriate endpoints and requirements for ecotoxicological representation when selecting test species.

The presented projects deal with the evaluation of invertebrate species sensitivity and the identification of species-specific hazards and risks within different scopes: The general comparability of sensitivities of communities living in different water types, and the risk assessment of an insecticide for re-registration. Both projects make use of SSDs, but only as part of a comprehensive view at existing data and information regarding the trophic level, the norm of reaction, the taxonomy, and the physiology and habitat ecology of the species. Additional evidence by further approaches such as mesocosm and monitoring data is also included to generate a consistent risk profile with low uncertainty, even when not exactly quantified.

2.2.2 Comparison of sensitivity distributions as check of representativity: Ecotoxicological sensitivity of groundwater organisms to pesticides

In recent years, the threat of chemical pollution to groundwater is no longer regarded as a concern for human drinking water resources only. With growing information on the complexity of the biological ground water community, the groundwater ecosystem is regarded as a good of protection in its own right. Risk assessment for aquatic systems so far is restricted to surface waters and is based on standard aquatic test results. The question is raised whether and in how far the common practice of risk assessment is also protective to groundwater communities. The major objective of the presented project for the German UBA (ref. 30) was the comparison of the sensitivity of physiologically / taxonomically comparable species from groundwater and surface water in acute toxicity studies with the same substances. Where possible, hints to chronic toxicity should be derived. Potential deviations of specific responses should be identified and interpreted in the light of properties specific for groundwater habitats. The investigations addressed microorganisms as well as metazoan species.

---

[2] www.eufram.com

2.2.2.1 Substances and methods

Microbiological investigations comprised functional tests (enzymatic activities) with a laboratory strain of the standard organism *Pseudomonas putida* and microbial communities of ground- and surface water samples, as well as of structural investigations of the microbial communities (genetic diversity) by RFLP and T-RFLP (by Thomas Lukow and Andrea Wenzel, Faunhofer IME).

Concerning the metazoan community, the sensitivities of groundwater organisms to three pesticides were investigated and compared with those of taxonomic relatives from surface waters in long-term toxicity tests. As crustaceans represent the most important groundwater taxa, copepod (Harpacticoida) and amphipod species were selected as well as Syncarida, a group only present in groundwater. The specimens were taken from groundwater bodies in the Ruhr and Main valleys. They were held and tested at about 10 °C in the dark. Breeding or the performance of chronic tests were not possible because of the life performances being retarded by a factor of 5–10 compared with surface water organisms. The sampling, selection, holding and testing of groundwater organisms was performed in cooperation with the crustacean specialists of the University of Oldenburg (Edgar Egert, Thomas Glatzel, Prof. Schminke).

The pesticides were selected according to their toxicity to crustaceans, to the availability of comparative data sets and mainly to the differences of modes of action. The fungicide cyprodinil inhibits anabolism, the pyrethroid lambda-cyhalothrin is extremely toxic to higher crustaceans due to the inhibition of the activity mediated by the axonic system (therefore not used in microbiological tests), and the herbicide bromoxynil-octanoate is mainly cytotoxic due to the non-polar narcotic action of the octanoate. The active ingredients (a.i.) were tested using commercially available single a.i. formulations; the concentrations were analytically controlled.

The effects were related to mean measured concentrations (lambda-cyhalothrin: isomers summed up), nominal concentrations (if analyses did not deviate by more than 20%), or initial concentrations (bromoxynil-octanoate because of fast dissipation and acute mode of action). The specimens were observed daily. Test duration was terminated by control mortality above 20%, the time until a stable effect concentration did not exceed 3–4 weeks (shorter duration with bromoxynil-octanoate).

Only lethal effects were assessed. The evaluation was based on the highest concentrations without mortality.

2.2.2.2 Results and discussion

The effects of the active ingredients cyprodinil and bromoxynil-octanoate were not measured to be more pronounced in microbial groundwater communities compared to the surface water communities. Additionally, the effects were retarded due to lower temperatures. Potential structural changes could not (yet) be observed by the methods used.

In lower crustaceans, cyprodinil caused lethal effects of the same sensitivity in surface water as in groundwater, when exposure duration was five- to tenfold higher. The groundwater amphipod *Niphargus fontanus*, even when exposed for the 10fold duration, was by two orders of magnitude less sensitive than the standard higher crustacean *Mysidopsis bahia*. Anabolic inhibition effects in groundwater organisms seem to be retarded 5-10 times compared to surface water species. Amphipods seem to be much more sensitive in surface waters, probably because of the extreme differences in metabolic turnover rates.

The distribution of sensitivity to lambda-cyhalothrin for surface water organisms covered the distributions of groundwater organisms after short-term exposure as well as after five- to tenfold exposure duration (Figure 3). Lower crustaceans from ground- and surface water exhibited similar sensitivity, whereas groundwater higher crustaceans showed lower sensitivity compared to surface water higher crustaceans, which have been tested to be most sensitive to pyrethroids. Syncarida seem to be as sensitive as surface water higher crustaceans when exposed for a five- to tenfold duration. The mainly acute effect of lambda-cyhalothrin is due to the nerve structure and activity, affecting lower crustaceans of ground- and surface waters in a similar way. Higher crustaceans in general exhibit more sensitive nerve structure and – in surface water – higher activity. Thus, groundwater higher crustaceans seem to be less sensitive. Syncarida, however, are highly adapted to the constant groundwater conditions and less flexible towards additional stress compared to *Niphargus fontanus*, resulting in retarded, but similar sensitivity as surface water higher crustaceans.

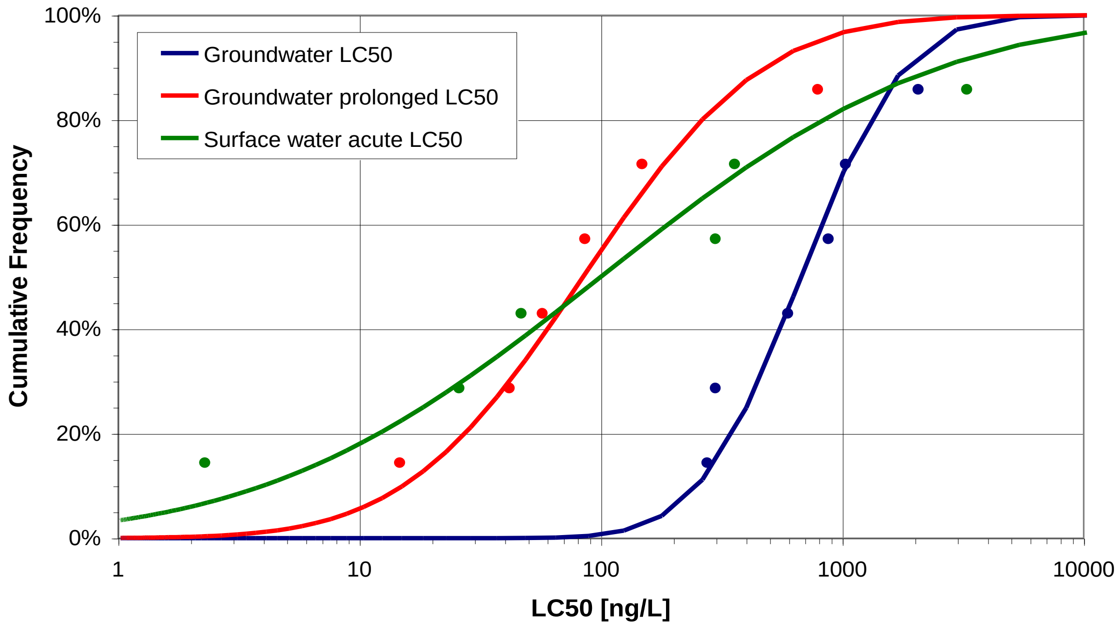


Figure 3: Species sensitivity distributions for ground- and surface water species to lambda-cyhalothrin

The slopes of the acute and prolonged groundwater sensitivity distributions were comparably steep indicating a low variability of results, whereas the variability of sensitivities in surface water was much higher, resulting in a much lower HC5. This an be explained by the by far more homogeneous habitat conditions in groundwater, having resulted in a more similar physiology due to evolutionary adaptation.

With bromoxynil-octanoate, hardly any differences in sensitivity between taxonomic groups or between groundwater and surface water organisms after short-term exposure could be observed. Representatives of Harpacticoida and Ostracoda had similar LC50s after 48 hours compared with the *D. magna* EC50; the development of the LC50 value over time was nearly identical in the representative of groundwater Syncarida and the standard higher crustacean *Mysidopsis bahia* (Figure 4). Because the substance dissipated quickly from the water even at groundwater temperatures, comparisons with static long-term exposure are neither possible nor necessary. Again, *Niphargus fontanus* tended to be less sensitive. Narcotic effects seem to be similar in ground- and surface water organisms after comparable exposure durations, irrespectively of the way of living.

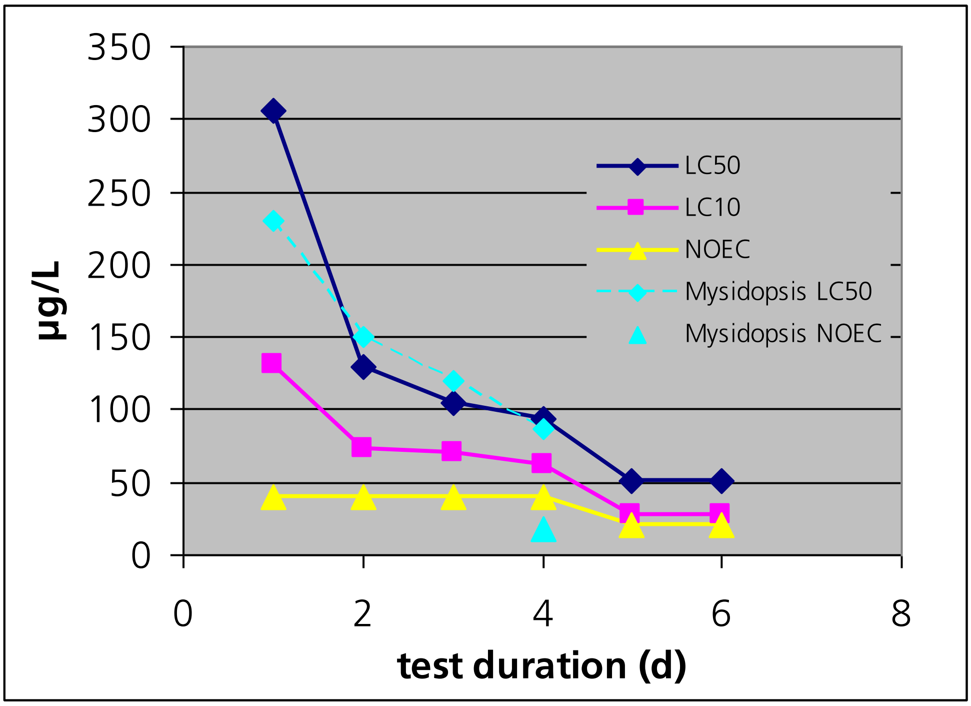


Figure 4: Effect concentrations of bromoxynil-octanoate to Syncarida

2.2.2.3 Conclusions

The sensitivity distribution of lethal toxicity to groundwater species can be represented by taxonomically and physiologically comparable species of surface waters. Inhibition of the anabolism or the activity affects groundwater organisms less or clearly retarded as compared to surface water organisms. No hints to a physiologically higher sensitivity of groundwater organisms could be observed. At the actual state of knowledge, risk assessment based on standard organisms seems to be sufficiently protective also for groundwater communities. If there are concerns about potential effects to higher crustaceans, a representative species (e.g. *Gammarus, Asellus, Hyalella*) should be tested in order to be protective regarding the limited norm of reaction of Syncarida (as shown with lambda-cyhalothrin). However, it has to be strongly emphasized that the population dynamics of groundwater species, if affected, allows only very slow and insufficient recovery by reproduction or recolonization.

2.2.3 Habitat specific sensitivity distributions: Use within pesticide risk assessment

2.2.3.1 Background and first approach

In 1956, carbaryl was introduced to the market as first carbamate-insecticide. Its mode of action is well described as inhibition of acetylcholinesterase (AChE). Brock et al. (2000) concluded from a literature review on single species and community level studies with AChE-inhibitors that the structure of aquatic communities is clearly more sensitive than functional characteristics of the ecosystem. The sensitivity of aquatic communities to direct effects was considered to be driven by the most susceptible taxonomic groups of crustaceans and aquatic insects. There is a comprehensive database on ecological effects of this acute highly toxic active ingredient, comprising an extraordinary number of tested freshwater invertebrate species (>30), several micro-/mesocosm studies and field observations. However, most of the toxicity studies are non-GLP studies with test durations between 24 and 96 h, partly without analytical verification of concentrations, and related to nominal or measured concentrations: EC/LC50-values vary between 1.7 and > 3000 µg/L (Figure 5).

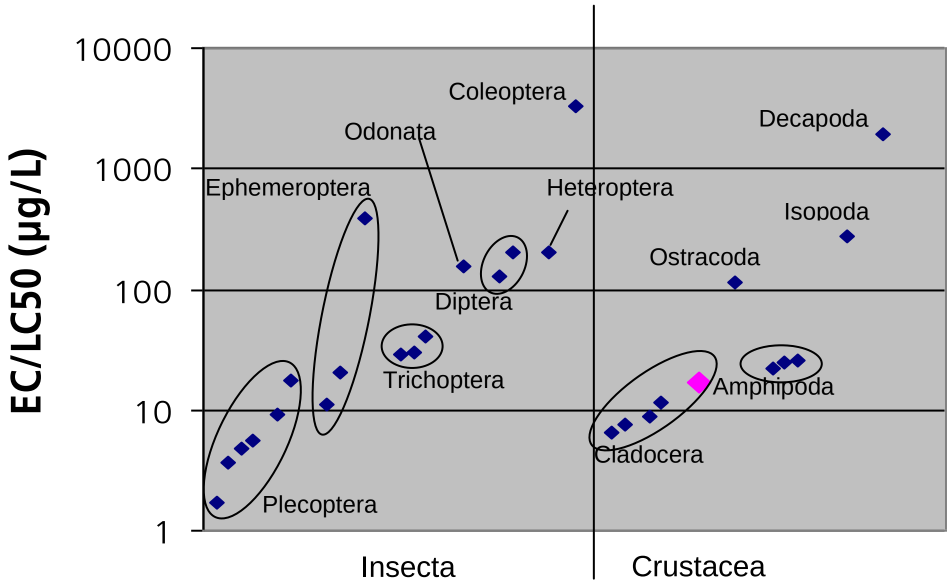


Figure 5: Distribution of acute toxicity of carbaryl to freshwater invertebrates (data from literature) with respect to taxonomy
Pink: GLP study with *D. magna,* the standard invertebrate organism

The presented work, sponsored by BayerCropScience, aimed at the derivation of a reliable “Ecologically Acceptable Concentration” (EAC) for freshwater invertebrates with uncertainty left as low as possible (ref. 65). The validity of the most relevant literature data was proved by performing well documented GLP studies with representatives of the most sensitive groups Plecoptera (*Isoperla grammatica*), Ephemeroptera (*Ephemera danica*), Cladocera (*Daphnia longispina, Chydorus sphaericus*), and Amphipoda (*Gammarus fossarum*). In addition, two mollusks as representatives of the important groups of Bivalvia and Gastropoda were tested, but these were shown to be just as insensitive as expected based on data for marine clams and freshwater gastropods from the literature, with acute LC50 values far beyond 1000 µg/L. The influence of introduced sediment on exposure of test organisms was also studied, as well as the effect of worst-case pulse exposure on the stonefly species. The most sensitive groups of organisms were identified. For these, the hazard potential was derived in relation to the potential duration of exposure.

2.2.3.2 Results and confirming work

The properties of carbaryl (water solubility, acute mode of action) permitted highly comparable study results, irrespective of study design, analytical verification of test concentrations, or the presence of sediment. A comparison of the toxicity data from our own GLP studies with the non-GLP data from literature revealed a notably high level of consistency: The tested species exhibited EC50/LC50 values within factors of 0.8 to 1.4 of the mean values of the appropriate taxonomic groups reported in the literature. Only in the case of the mayflies was the verification of literature data not possible because stream mayflies (*Cynigma*, *Ameletus*) appeared to be much more sensitive than stagnant water mayflies (*Cloeon*). *Ephemera danica* was found to be of medium sensitivity. Taking into account the habitats of *E. danica* (which digs in fine substrates of calm habitats in midland streams), a comparison of our GLP data with means of literature data (species living in comparable habitats) shows consistent results.

The sensitivity of the tested species seems to be dominated by eco-physiological properties. It is hypothesized that the water exchange rate at respiratory membranes drives uptake and internal exposure. The three most susceptible groups of organisms (Figure 6) are

1) Insect larvae from fast-flow habitats (stoneflies and heptageniid mayflies, seven species, LC50 1.7 – 17 µg/L)
2) Cladocerans as planktonic high-performance filter feeders (six species, EC50 6.4 – 17 µg/L)
3) Insect larvae and crustaceans from moderate-flow habitats (gammarids, mayflies, caddisflies, eight species, LC50 20-41 µg/L)

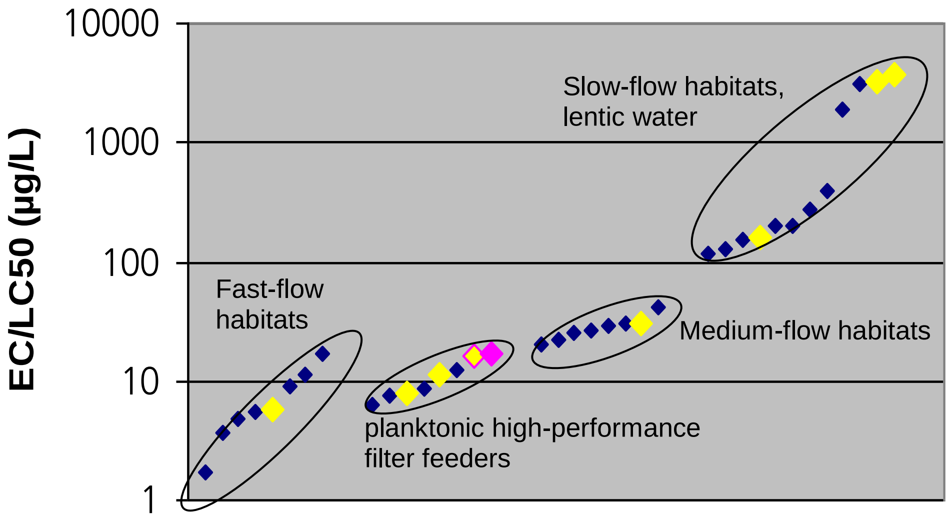


Figure 6: Distribution of acute toxicity of carbaryl to freshwater invertebrates with respect to habitat ecology
Pink: *Daphnia magna*, the standard invertebrate test organism, yellow: our GLP studies

These susceptible ecological groups were examined with reference to general principles of sensitivity and risk, the sensitivity of available effect data, hypotheses of risk under field conditions, and study results helping to reduce remaining uncertainty.

1) Insect larvae from fast-flow habitats

Organisms adapted to habitats dominated by fast water flow usually expend a lot of energy simply maintaining their position (e.g. trout swimming against the current, stonefly larvae or heptageniid mayfly larvae holding fast to stones). Oxygen is not limiting in such habitats. Organisms adapted to these conditions do not usually need to move actively gills and water, but to provide surface membranes with high uptake potential and to align them to the current. If they enter habitats of low current velocity (e.g. static laboratory testing conditions), they are forced to move in order to generate sufficient flow at their gill membranes thus permitting sufficient oxygen uptake.

Because of the contact probability, these organisms tend also to absorb larger quantities of substances via the gills. This may be the main reason for the toxicological sensitivity of salmonids (which are frequently observed to be higher than that of other fish) and may accordingly explain why stonefly larvae are much more sensitive than other aquatic arthropods, a fact also shown by the reference data. This hypothesis is supported by the fact that the mayfly larvae of the Heptageniidae, which inhabit sites of high current velocity, were shown to be as sensitive to carbaryl as the most sensitive stoneflies, whereas *Ephemera danica*, mayfly larvae inhabiting fine sediments of calm microhabitats of midland streams, are far less sensitive. The LC50 of 5.8 µg/L exhibited by *Isoperla grammatica* is typical of the most sensitive data reported for wild life stoneflies in literature with a mean LC50 of 7.0 µg/L.

Sites of high current velocity are characterized by pulse exposure followed by fast removal of pesticides by unloaded water. Thus, the high flow rate causing adaptations which appear to be responsible for high uptake (high doses), also results in reducing the duration of exposure. When stonefly larvae were exposed for only one hour, which can be regarded as a worst-case situation after a pesticide loading, no effect occurred even at 20 µg/L. At higher concentrations, a knockdown effect occurred starting after 20-30 min, followed by recovery of all individuals up to the highest concentration of 100 µg/L. This is consistent with the findings in the literature, indicating no significant mortality in heptageniid mayfly larvae (*Cynigma sp.*) after 96 h following one hour carbaryl exposure at 102 µg/L. The LC50 was determined to be 165 µg/L. Concentrations above 20 µg/L, causing knock-down effects, cannot be regarded as safe, because the affected individuals will drift to downstream habitats to which they might be less well adapted, or they could be predated.

The available data show that a concentration of 20 µg/L initially generated in a habitat with high current velocity can be regarded as safe for a typical organism adapted to that habitat. This has possibly been confirmed by a field study, measuring peak concentrations of carbaryl after application and the impact on the aquatic community of a stream in a drought and a non-drought year: Depending on the water flow, the peak concentrations, occurring within two hours after the applications, were 85.1 and 12.6 µg/L, respectively. At a peak concentration of 12.6 µg/L, no difference in organism drift compared to an uncontaminated upstream reference site could be observed. At a peak concentration of 85.1 µg/L, drift was more variable and could be attributed to increased drift of heptageniid mayfly (as the only taxon affected) within three hours after pesticide application. No stoneflies were present.

2) Cladocerans as planktonic high-performance filter feeders

Cladocerans (daphnids) are mostly filter feeders of lentic waters. When using their phyllopod legs they filter high amounts of water resulting in high contact probability of the gills and the gastro-intestinal part with dissolved or dispersed substances. Daphnids belong to one of the taxonomic groups shown to be highly sensitive to carbaryl. In order to reduce uncertainty about the variability of daphnid sensitivity, three species were tested: a laboratory strain of *D. magna* as standard test organism, and two species indigenous to ponds and small lakes (*D. longispina* and *Chydorus sphaericus*), representing Cladocerans of different habitats. The EC50 of 7.8 µg/L for *D. longispina* is fully typical of the most sensitive data reported for wild life daphnids in literature ranging from 6.4 to 11.5 µg/L. Considering chronic effects including reproduction, a *D. magna* reproduction study showed 20% mortality at a mean concentration of 8.0 µg/L, whereas neither mortality nor effects on reproductive performance were observed at 7.1 µg/L. Thus, the effect pattern is clearly dominated by lethal effects.

All cladoceran species are mainly parthenogenetic with a high population growth rate, and make use of sexual reproduction only under the influence of environmental stress. They are able to form ephippiae (resting eggs), which remain in the topmost layer of the sediment and which are the source for re-population after extinction of the parental generation. Thus, Cladocera are able to recover from severe population reductions in various ways: If reduced to a few individuals, they may recover by rapid population growth. If extinguished, they may recover from within the system by growth from permanent eggs, or they may be re-introduced by ducks or other water birds. The worst-case scenario of an isolated pond must focus on recovery from within the system. Such recoveries have been demonstrated for different test substances when the test substance concentrations had fallen below a critical level.

For derivation of an EAC, the fate of the test substance as well as the recovery potential of the organisms at acute risk has to be taken into account. For carbaryl, with a half-life of well under 2 days in the water column of realistic water/sediment systems, there is no reason to expect an observable effect on *D. longispina* over a prolonged period at concentrations near to the laboratory EC50 of approximately 8 µg/L. This is confirmed by a publication demonstrating the relationship between the age of daphnids at test start and the sensitivity to Carbaryl. The lowest 48 h EC50 of approximately 7 µg/L was reported for 1 d old individuals. 3 d, 5 d and 10 d old individuals exhibited an EC50 of approximately 14 µg/L, 18 µg/L and 26 µg/L, respectively. In a second test series, 45 % of the most sensitive stage between 12 and 24 h of life survived 10 µg carbaryl/L. The differences in sensitivity were attributed to the physiological activity and especially to the state of molting. The manifestation of effects on the field population level is supposed to depend on the age structure of the population. This was investigated by a further paper exposing *Daphnia pulex* populations at different dynamic phases (growing phase with a high proportion of young individuals, density peak with a medium proportion of young individuals, declining phase with a low proportion of young individuals) to 15 µg carbaryl/L, which was supposed to kill only juvenile life stages. The control oscillation took 6.3 days from the minimum density after the first peak to the second peak. An application during the growing phase before the first density peak caused a reduction to 15% and a recovery within 9.6 days, resulting in similar second peak densities in a similar time interval compared to the controls. An application during the decline phase after the first peak reduced the minimum density to only 6% of the peak density, but enabled full recovery within 8 days. The most lasting effects were caused by an exposure during the peak density phase. In this phase, competitive stress is maximized and food resources are minimized. The population decreased to 0.7% of the peak density and needed 17.6 days for full recovery.

A number of mesocosm studies on zooplankton with carbaryl are available, demonstrating effect concentrations as well as re-population with cladoceran species after partial and total extinction of daphnids. The most sensitive data are short-term effects (4 d) on the biomass of *D. galeata* biomass of about 50% at an initial concentration of 5 µg/L, and by definitely more than 50% at 7 and 10 µg/L. Because surviving individuals were found at these concentrations (no survivors at 20 µg/L), Brock et al. (2000) evaluated the effects up to 10 µg/L as slight effects assumed to be followed by rapid recovery. There is even evidence from other studies to postulate a complete recovery from ephippiae within two months at concentrations higher than the EC50 by one order of magnitude, reporting full recovery at initially 100 µg/L of carbaryl (lowest concentration) after six weeks.

However, as an EAC of 10 µg/L would be based on a substantial initial effect, independently from the model ecosystem and population studies the recovery potential of cladocerans was estimated by means of a simple population model by Udo Hommen. Using worst-case assumptions about the intrinsic growth of increase (0.17 /d for *Bosmina longirostris*) and sensitivity (EC50 of 6.4 µg/L for *D. pulex*), the predicted time to recover from a pulse exposure of 10 µg carbaryl/L back to 95% of the carrying capacity was predicted to be 18 days. Considering uncertainty about sensitivity and population growth, the mean recovery was estimated to be within 8 days. For the evaluation of a proposed EAC of 10 µg/L, the following conclusions can be drawn:

- The 48h acute immobilization tests with cladocerans represent absolute worst case situations of exposure as well as of effects, as 0-24h old individuals are exposed which were shown to be clearly the most sensitive stages. In field populations, an exposure to 10 µg carbaryl/L will lead to the loss of only a limited part of the population.
- All effects up to a reduction by 99% are able to recover within 3 weeks. 10 µg carbaryl/L will most probably cause population losses of between 50 and 80%. Reductions by 80% correspond to the normal oscillations and recover within a week.
- The most severe effects were shown to hit a population reaching the habitat capacity. However, full recovery within three weeks was even possible after an exposure of a very sensitive species to 15 µg carbaryl/L.

3) Insect larvae and crustaceans from moderate-flow habitats

Whereas risk-reducing mechanisms can be identified for the two most sensitive groups (i.e. cladocerans as planktonic filter feeders and stoneflies as representatives of fast flow sites), this is more difficult for other sensitive organisms: species without the reproductive recovery potential of daphnids and without the short pulse exposure situation of fast flow habitats. The available data demonstrate that slow-flow site macro-invertebrates are usually not very sensitive, especially when inhabiting stagnant waters. A very homogeneous group is formed by the organisms shown to be the most susceptible species left: organisms of moderate-flow habitats represented by amphipods, caddisflies, and some mayflies. The representative taxon tested was the amphipod gammarids: LC50 data of the most abundant North American and European gammarid species are included in the data set, showing very close LC50 values irrespective of nominal or measured values. The data are highly consistent and thus reduce the level of uncertainty. The potential reduction in the time and level of exposure by water flow offers additional certainty for assessment of effects in the field. For all gammarid species, the behavioral reduction of exposure by their well-known migratory properties can be regarded as an additional safety buffer, provided access is available to resource habitats for recolonization.

As it cannot be assumed that the species of the group in question are present only in moderate-flow habitats, the argument of dilution and the postulation of reduced risk by pulse exposure as used with stoneflies and heptageniid mayflies are not applicable. The assessment of risk is thus worked out by means of an SSD of acute toxicity data of the community inhabiting moderate-flow habitats. In order to include additional safety, the “concentration protective for 95% of the species” (HC5) was calculated based on NOEC data (Figure 7). Since acute NOEC data were not available for most of the tests, the LC50s were reduced to virtual NOECs by assuming the same slope of the concentration-effect relationship as for *G. fossarum*, which seemed to be justified by the unique mode of action: the factors between EC/LC50 and NOECs from the GLP studies all ranged between 1.41 and 1.70.

The results of the HC5 calculations show that gammarids are well representative of moderate-flow habitat species: the 90% confidence intervals of the HC5 for all moderate-flow species and that for gammarids only were nearly identical.

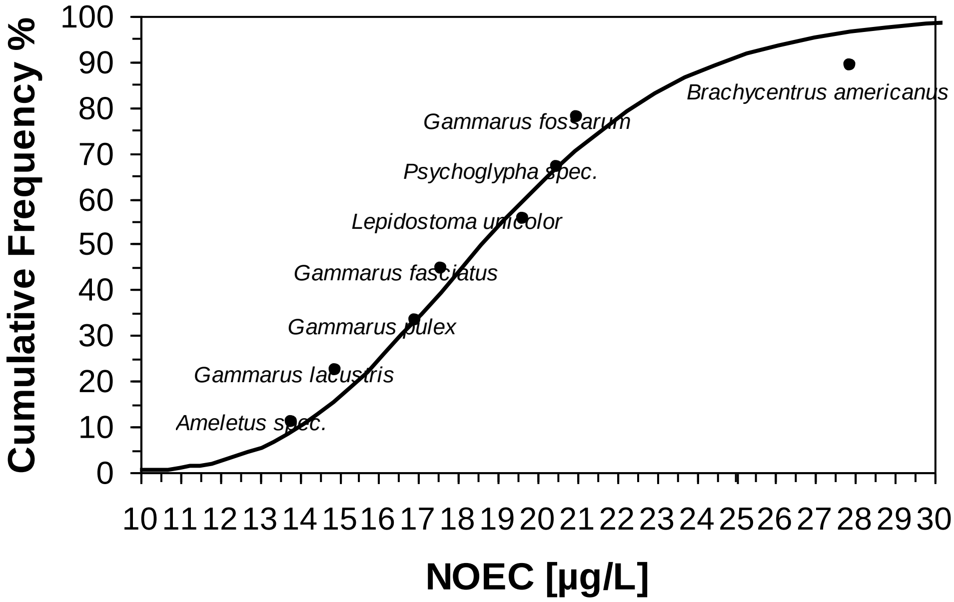


Figure 7: Cumulative frequencies of the species sensitivity distribution to carbaryl of moderate-flow habitat species
The curve is based on mean and standard deviations of the log-transformed NOECs.

2.2.3.3 Result summary and conclusion

The sensitivity of the tested species to carbaryl seems to be dominated by eco-physiological properties. It is hypothesized that the water exchange rate at respiratory membranes drives uptake and internal exposure. The three most susceptible groups of organisms are

1) **Insect larvae from fast-flow habitats** (stoneflies and heptageniid mayflies, 7 species, LC50 1.7 – 17 µg/L). After pulse exposure for 1 h to a representative stonefly species (LC50 at 96h exposure: 5.8 µg/L), a NOEC of **20 µg/L** was observed after 96 h. At 100 µg/L, the initial knock-down-effect occurred not before 30 min. This was consistent with pulse exposure data with carbaryl and heptageniid mayflies.

2) **Cladocerans as planktonic high-performance filter feeders** (6 species, EC50 6.4 – 17 µg/L). The EC80 of a sensitive species was determined with **10 µg/L**. That effect level was shown in micro-/mesocosm studies, population studies and model calculations to recover completely within 3 weeks. At 100 µg/L, a recovery from permanent eggs (ephippiae) was demonstrated after 6 weeks by mesocosm studies from the literature.

3) **Insect larvae and crustaceans from moderate-flow habitats** (gammarids, mayflies, caddisflies, 8 species, LC50 20-41 µg/L). For this group it is not feasible to suppose obligatory short-term exposure (see 1)) or rapid recovery by reproduction (see 2)). Therefore, a NOEC-based SSD was calculated. The 95 percentile of the NOEC is **12.7 µg/L**.

Based on the most sensitive group represented by planktonic feeders (daphnids) an overall EAC of 10 µg carbaryl/L for all freshwater invertebrates is proposed. This EAC is considered conservative taking into consideration additional safety buffers under wildlife conditions such as rapid degradation of carbaryl, the particular re-population capacity of cladocerans via ephippiae, and that of gammarids and insects via immigration, if a connection to resource habitats for recolonization is given.

As a most recent development in the risk assessment of carbaryl, the sensitivity of shrimps was included in the risk assessment. The marine shrimp *Mysidopsis bahia* was not included in the first assessments. However, is regarded representative for freshwater shrimps like *Palaemonetes spec,* which lives in lakes in southern Europe. The LC50s from literature and GLP studies for both species range from 5 to 7 µg/L. These results represent the low part of the planktonic filter feeder group and thus are consistent with the presented concept, but the recovery potential is clearly lower and can not be extrapolated from cladoceran species. Consequently, a concentration of 10 µg/L is not protective for freshwater shrimps. However, as appropriate habitats are deep waters, risk mitigation measures reducing initial carbaryl concentrations to 10 µg/L in ditches of 30 cm depth are also protective for freshwater shrimps.

2.2.4 Implications for future work

The hypothesis supported by the carbaryl example (species sensitivity is mainly due to uptake and internal exposure, depending on the water exchange probability at respiratory membranes and thus being highly correlated with oxygen demand) is of general importance for ubiquitous modes of action and needs validation by uptake experiments with an appropriate set of organisms. These experiments should be performed with a ^{14}C labeled test substance (preferably carbaryl). The results can be used for calculating uptake rates and check them for taxonomic or habitat specificity. A data set representative of the main taxonomic groups or habitats can be used to approach a dose-related risk assessment rather than restricting it to concentrations in water.

The presented examples demonstrate that including ecological information and additional evidence by different approaches can substantially reduce uncertainty. For meeting the requirements of higher tier risk ecological assessment, SSDs should be used in a two-step-approach:

a) Species of all taxonomic groups being at risk should be included. This first step should ensure the identification of sensitive groups with low uncertainty. All information on effect concentrations exposure and hazards should be considered when selecting test species. The slope of the distribution will be flat, the fit may not be satisfactory (as shown for carbaryl), the HC5 will be very low and uncertain, and thus of limited value.

b) The most sensitive species should be clustered according to taxonomy / physiology and / or habitat ecology. Some clusters may be excluded from further assessment due to obvious insensitivity or missing exposure. The risk assessment for the remaining clusters can either be performed separately by including further approaches (as shown for carbaryl, but only possible with as much information) or by calculating an SSD of combined species. The slope will be steeper, the HC5 will be higher, and the uncertainty of the assessment will be lower.

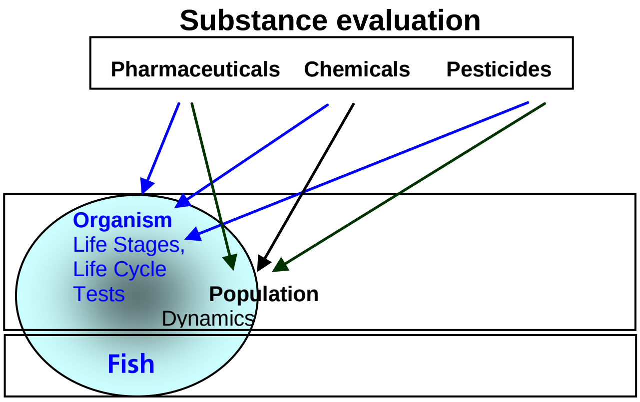
Nevertheless, lowest toxicity values near or even below the HC5 should be evaluated for the ecological severity of statistical endpoint. Representative chronic risks should be accounted for. If the SSD is based on EC50s instead of NOECs or EC10s, habitat exposure and recovery potential should be considered. This will reduce substantial uncertainty of the overall assessment beside the reduction of statistical uncertainty.

2.3 Sensitive life stages and performances: Fish

2.3.1 Introduction

When aiming at the protection of ecosystems from risks by chemical pollution, the primary focus is on the protection of populations rather than individuals, which are in the focus when aiming at the protection of human health. Chronic tests should allow the evaluation of all endpoints of relevance for population growth and maintenance:

- Stage specific survival rates
- Time to first reproduction
- Offspring number per female
- Sex ratio



These endpoints are most important for the hazard- and risk assessment and can be extrapolated to populations by modeling the intrinsic rate of population increase. Up to now, there are no generally accepted and standardized test protocols for fish covering the mentioned endpoints besides the testing of early life stage toxicity (only part of stage specific survival rates) and juvenile growth (only part of endpoint “time to first reproduction”).

I have worked on the interaction between toxicity and fish population ecology since my diploma thesis, including the generation of data for the development of a stochastic zebrafish population model and its application to the guppy (*Poecilia reticulata*), a species with a different reproductive strategy (ref. 1), 2), 4)-9) , 33), 34), 37)). A major part was the performance of a two generation test (2-GT) with the guppy and full life cycle tests (FLCT) with zebrafish (*Danio rerio*), contributing to the development of the first zebrafish FLCT protocol proposal by Nagel (updated version: Nagel, 1998). This work was of limited international importance as it was regarded as academic and not suited for regulatory purposes. This view changed considerably with the identification of endocrine disrupting chemicals (EDCs) in the environment.

EDCs can interact with endogen hormone receptors or active centrums of enzymes. As they are active within cells, they have properties, which particularly constitute a risk for aquatic communities, especially fish (see chapter 1.2; (ref. 49)). Due to the evolutionary conservatism of the steroid system, fish susceptibility is close to that of mammals. Thus, knowledge on mammalian toxicology generated to protect human health can also be used for assessing the hazard to fish. At the same time, the fish system, if it can be tested more comprehensive in due time, could be used for a first assessment of the hazard potential of EDCs to mammals. In the following, the definition EDCs is restricted to substances, which demonstrated to disrupt specifically endocrine mediated life performances, such as development and reproduction. A simplified mechanistic classification of direct sexual endocrine effectors in fish results in the two main groups of receptor interactions (i.e. agonists or antagonists of the estrogen or androgen receptor) and inhibitors of enzymes of the hormone synthesis. The latter can be divided in inhibitors of the testosterone synthesis (general inhibition of sexual steroids), the 11-hydroxylase (inhibition of the synthesis of 11-keto-testosterone (11-kT), the active male fish hormone), and / or the aromatase (inhibition of the 17-beta-estradiol synthesis). As a result, effects described as estrogenic, androgenic, anti-estrogenic or anti-androgenic may occur as well as various mixtures of them (ref. 49). In the following, projects of the last ten years are presented, focusing on the development of fish FLCT designs/protocols for different objectives. The work contributed substantially to the development of a test protocol for full life cycle tests including endpoints being indicative of endocrine disruption and relevant for population dynamics (key publication: Wenzel and Schäfers, 2001 (ref. 53)). It also contributed to the derivation of a fish testing strategy for

EDCs, highlighting the predictive potential of biomarkers (summarizing key publication: Knacker et al., 2010 (ref. 27)). Matthias Teigeler performed a main part of the experimental work during his PhD thesis (ref. 3).

With 22 fish full life cycle, two- or multi-generation tests in public or confidential GLP projects, of which 17 tests were performed with potential or definitive endocrine disruptors, the IME laboratory has long experience and high expertise concerning the development, validation, investigation and assessment of indicative and population relevant endpoints.. This was acknowledged by the nomination as national expert for the OECD fish drafting group. The IME laboratory was the lead laboratory for zebrafish during the validation of the 21d fish assay (OECD TG 230), took part with two species (medaka and zebrafish) in the validation of the fish sexual development test FSDT (OECD 234) and contributes to the development of OECD definitive fish test protocols. I participated in the setup of the OECD fish testing framework (finalized in 2011) and was asked for expert opinions by the German IVA (ref. 60) and BVL (ref. 63). As invited ecotoxicological expert I participated in the CEFIC workshops 2008 on fish life cycle tests in Palma di Mallorca (ref. 26) and on fish gonad histopathology in Ermatingen as well as in the ECETOC workshops on endocrine disruption in Barcelona 2009 and Florence 2011.

2.3.2 Fish Full Life Cycle Tests with estrogenic substances: natural substances, pharmaceuticals, chemicals

The investigation of endocrine effects on fish by anthropogenic substances was mainly focused on the analysis of biomarkers in environmental samples or on *in vitro*-screening tests for substance evaluation. In both cases, the ecological relevance of the investigated endpoints is unclear. This is also true for the measured environmental concentrations of substances of concern. Within two EC projects, one 5th framework R&D-project led by Helmut Segner, Environmental Research Centre UFZ in Leipzig, Germany, and one tender contract, zebrafish FLCTs were performed to identify threshold concentrations for effects being relevant for fish populations. These effect concentrations were correlated with histological effect patterns and biomarker responses (ref. 14)-16), 18), 38), 40), 44), 46), 53)).

2.3.2.1 Substances and methods

The investigated substances were ethinylestradiol (EE2) as most important contraceptive estrogen and strong positive control, p-tert-octylphenol (OP) as clearly defined alkylphenol, genistein (GEN) as phytohormone with high estrogenic potential found in *in vitro*-tests, and bisphenol A (BPA) as important industrial chemical and weak estrogen.

The two-generation test (EE2) or full life cycle tests (EE2 at 2 ng/L, OP, GEN, BPA) were conducted under flow-through conditions (EE2 and OP) or semi-static with renewal of the test solution three times weekly (GEN and BPA). The tests were started with 100 fertilized eggs in two replicates for each concentration and the untreated control. The eggs were bred under exposure until sexual maturity and reproduction. The early life stages of the following generation were also exposed. The investigated endpoints were hatch, stage-specific mortality, growth (photographical evaluation of lengths), time until spawning, number of eggs per female and day, and fertilization rate. The test concentrations were analyzed by GC-MS/MS at least weekly. After test termination, gonadal histology and measurements of the yolk precursor protein vitellogenin (VTG) in blood plasma of the adult fish were performed by our UFZ project partners Gerd Maack and Martina Fenske, respectively.

2.3.2.2 Results

The analytically measured test concentrations for EE2 were 0.05, 0.3, 1.1, 2.0 und 10 ng/L. For OP, they did not deviate by more than 20% from the nominal concentrations of 1.2, 3.8, 12 und 38 µg/L. For genistein, mean measured concentrations were 0.45, 1.3, 4.2, 20 and

50 µg/L. For BPA the mean measured initial concentrations three times weekly did not deviate by more than 20% from the nominal concentrations of 94, 187, 375, 750 und 1500 µg/L, whereas the geometric means of fresh an old mean measured concentration were 11, 24, 40, 86, and 157 µg/L.

Regarding mortality of early life stages there was no consistent concentration-effect relationship for EE2, OP und BPA. GEN caused massive effects at the two highest concentrations: At 50 µg/L all embryos dies during the second day when starting movements, at 20 µg/L they died during hatch. This can be explained by an inhibition of protein kinases, which is described as specific effect of genistein and is not specifically estrogenic.

The result data for the endpoints juvenile growth, time to first spawning, number of spawned eggs and fertilization rate are correlated in all tests. Estrogenic substances affect gonadal development: Untreated male zebrafish develop testes after having run through a phase of an undifferentiated ovary (protogyn gonad development, Takahashi, 1977). This process seems to be retarded by the influence of estrogenic acting substances, resulting in reduced juvenile growth, a kind of arrest in the protogyn phase and a prolongation of the time until start of the male mating behavior. This clearly observed effect led to a lack of the cue for female spawning and a retardation of reproduction (Figure 8).

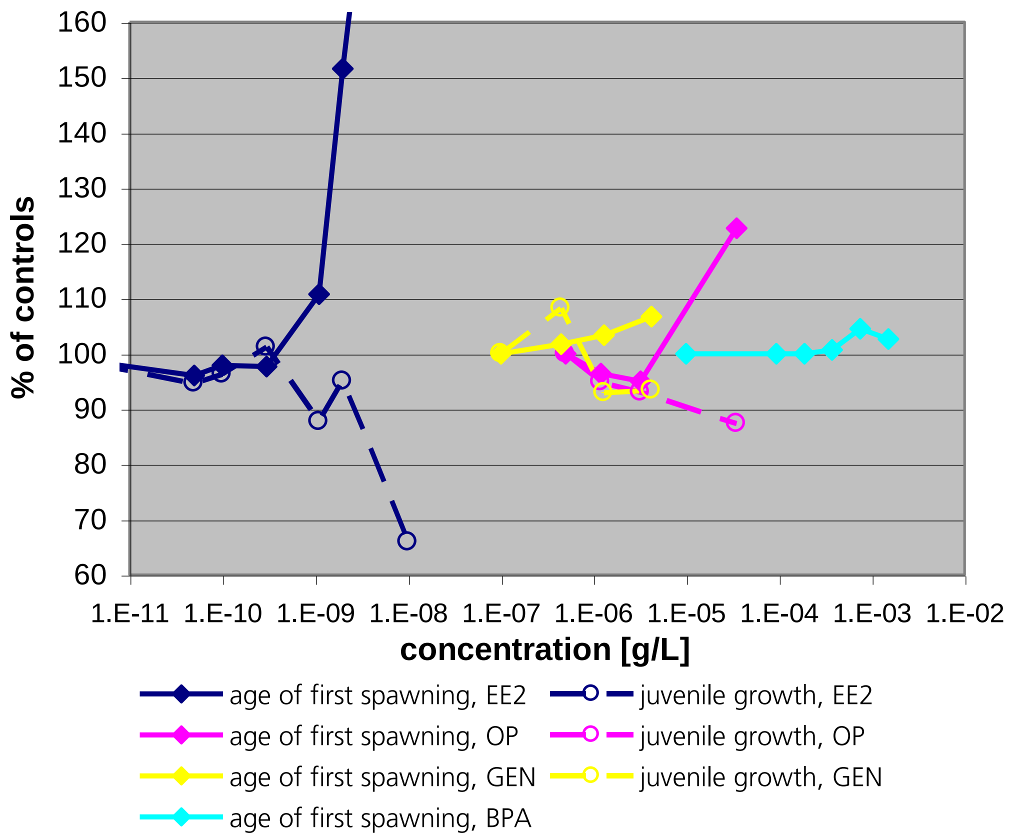


Figure 8: Growth and sexual development parameters in zebrafish FLCTs with estrogenic acting substances
Ethinylestradiol (EE2), p-tert-octylphenol (OP), genistein (GEN), bisphenol A (BPA); age of first spawning and juvenile growth (length) given as % of controls

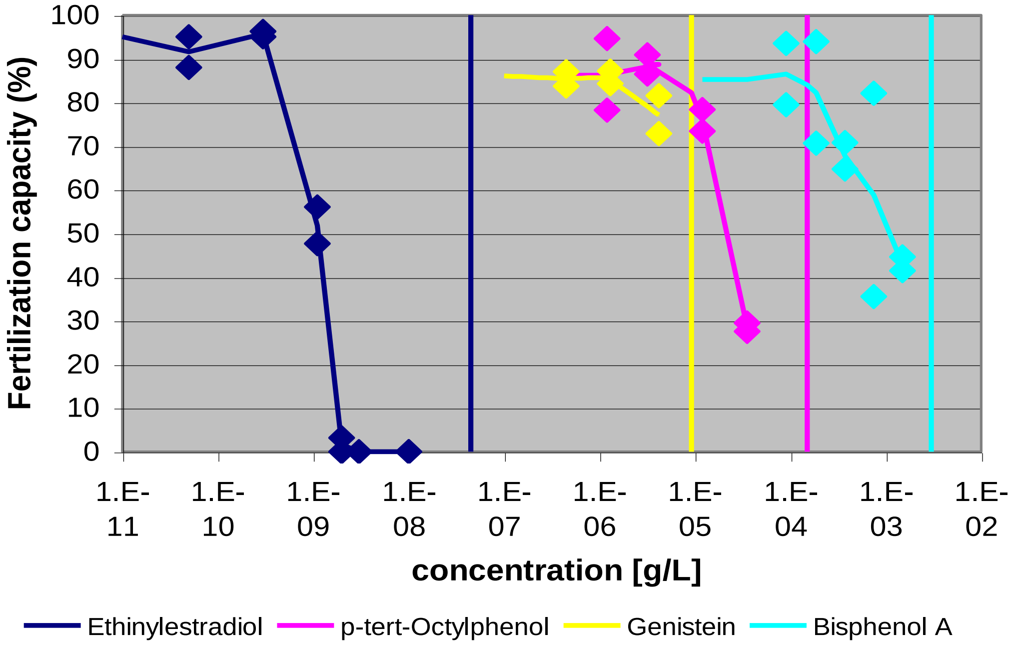


Figure 9: Fertilization capacity as most sensitive and relevant estrogenic effect in zebrafish FLCTs and lethal effects (LC10) (vertical lines)

After that retardation or, at 10 µg/L of EE2, after three months recovery in untreated water, male mating behavior started. Within short time, no difference to controls could be observed any longer regarding vigor of mating or number of spawned eggs. However, the fertilization rates were and remained clearly reduced (Figure 9). Thus, due to its high sensitivity, the maintenance of effect and the ecological consequences of spawning unfertilized eggs, the fertilization capacity was identified as most relevant endpoint for estrogenic effects on fish populations. This was confirmed by Young et al. (2002), who reviewed all available life cycle test information with EE2 to derive environmental quality standards for the British Environment Agency. As the fertilization rate had not been included in the other studies, they were less sensitive or regarded less relevant and Young et al. (2002) based their assessment on the presented Fraunhofer IME results.

The retardation of juvenile growth, the prolongation of time until first mating and spawning and the reduction of the fertilization rate can be characterized as estrogenic syndrome in zebrafish. It is clearly exhibited at exposure to EE2, less sensitive but also clear at exposure to OP and least clear at exposure to the weak estrogen BPA, which significantly affected only the fertilization rate. When exposed to GEN, the effect is nearly completely overridden by the lethal effects on the early life stages.

2.3.2.3 Conclusions

The hazard potential of estrogenic effects identified for OP and BPA is covered by the common uncertainty factors applied to acute LC50 data (Table 2). The estrogenic threshold value for GEN would be covered by results of an early life stage tests and an uncertainty factor of 10. The rapid biodegradation of GEN and BPA further reduces the risk. For EE2, however, estrogenic effect threshold concentrations are far below effect concentrations determined in standard tests. They are in the upper range of measured environmental concentrations without applying any uncertainty factor (ref. 40). Based on the IME results, Young et al. (2002) derived a PNEC at 0.4 ng/L.

Table 2: Effect concentrations and effect ratios of different estrogenic substances

	Ethinyl-estradiol (EE2)	**Genistein (GEN)**	**p-tert-octylphenol (OP)**	**Bisphenol A (BPA)**		**Relative potency related to EE2, molar concentrations**		
	µg/L	**µg/L**	**µg/L**	**µg/L**		**GEN**	**OP**	**BPA initial**
Lethal effects								
LC_{50} 96h (acute)	1 700	>1 900°	370	4 600#		>1.2	0.31	3.5*
LC_{50} 28d	100						5 300*	60 000*
LC_{50} larval toxicity		11.2				120*		
Fertilization rate				initial	total			
EC_{50}	0.0011	-	28	1 450	165		37 000	1 700 000
EC_{10}	0.0006	-	13.5	390	44		33 000	900 000
LOEC	0.0011	4.2	35	1 500	157	4 200	49 500	1 800 000
NOEC	0.0003	1.3	12	750	86	4 800	58 000	3 300 000
Acute/chronic	**ratio**	**ratio**	**ratio**	**ratio**				
LC_{50} 96h / EC_{50} fertility	1550 000	> 450**	13	3.2	28			
LC_{50}96h / NOEC fertility	5700 000	> 1460	31	6.1	53			

° > water solubility
LC50 for
* most sensitive lethal data compared;
** for LOEC fertilization rate

2.3.3 Considerations for a definitive fish test protocol for endocrine hazard assessment

2.3.3.1 Modifications of the zebrafish full life cycle test protocol

The test protocol developed by the Mainz working group of Roland Nagel (1998) intended to cover effects on stage-specific survival, growth and reproduction. To achieve sufficient statistical power for the variable fecundity endpoint, at beginning maturity of fish it was tried to identify the developing morphological sexes of fish and to set up groups of equal composition for the reproduction phase. Nagel (1986) proposed a composition of six females and twelve males to provide sufficient mating pressure for high fecundity and fertilization rates.

However, this approach has some severe disadvantages. Generally, a selection of only six females is a striking bias to the most clearly developed female fish, focusing on a potentially less sensitive percentile of the normal distribution. As the not selected fish often could not be reared under same flow-through exposure conditions as the selected groups, a re-grouping of reproductive groups by exchange of individuals was associated with additional variability. The grouping process took considerably time and prolonged the overall test duration. With respect to endocrine disruption, the test protocol did not include the determination of the time to first spawning and the sex ratio.

These shortcomings were identified during the EC projects on estrogenic effects (see 2.3.2) and GLP studies (ref. 43). Consequently, the protocol was modified.

- At the end of the early life stage phase (starting with 100 fertilized eggs per replicate), the fish are randomly reduced to 50 fish per replicate to achieve equal conditions for juvenile growth, which is highly density-dependent.

- After 9-10 weeks, the fish groups are randomly reduced to 30 fish to achieve equal conditions for reproduction; spawning trays are included[2]. There is no selection of sexes. The number of 30 is considered high enough to guarantee sufficient fish of both sexes. As zebrafish mate and spawn by variable formation of pairs in the group and without male dominance behavior, the approach represents low disturbance. Own investigations on the effect of different sex composition on reproductive output revealed, that the optimum composition is around 50:50, but relatively robust to shifts (ref. 56) (Figure 10).

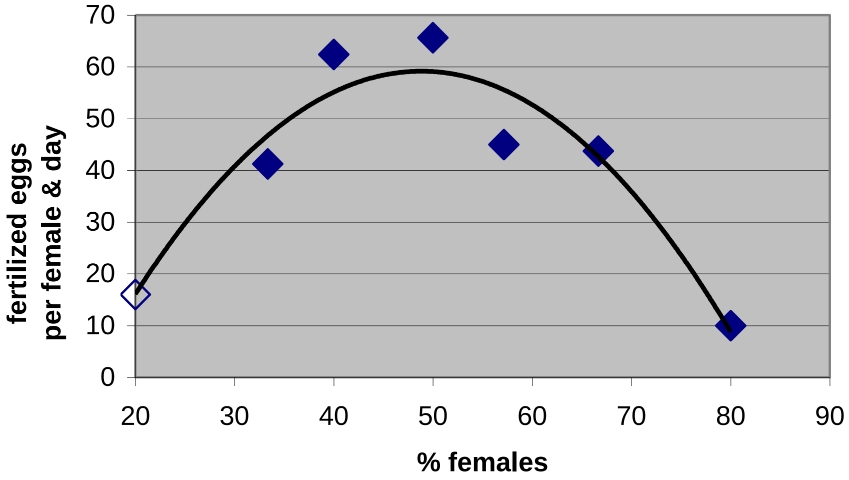


Figure 10: Reproductive output depending on sex composition of spawning groups

- At the end of the investigation of reproduction and after the filial generation has successfully passed day 21 of life, the parental groups are sacrificed, individual lengths and weights determined (and for endocrine hazard assessment blood sampled by cardiac puncture for VTG and 11-kT analysis). The fish are opened and the gonads are inspected for sex, dissected and prepared for histological investigation. The number of females per group is used to recalculate the number of spawned eggs per female. Thus, the endpoints time to first spawning, fertility, fecundity (eggs per female), and sex ratio (percentage of females, males and undeveloped) can be differentiated and added by the integrative parameter of fertilized eggs per replicate.

These modifications were discussed and implemented in a ring-test sponsored by the German UBA (ref. 57).

2.3.3.2 Two-generation tests

As member of the OECD fish drafting group on endocrine effects, nominated by the national coordinator (UBA), I participated in the international discussions on definitive fish testing, which is lead by the USA, Japan and Germany. The US presented a draft protocol on a fish two-generation test (2-GT) (OECD 2003). The proposed test starts with the exposure of adult fish and the investigation of reproduction, follows exposure of the filial F1 generation from embryo to the reproductive stage and ends with the exposure of the early life stages o the F2 generation. The protocol mainly focuses on fathead minnow as test species. At the IME, it was adapted to zebrafish by including the approaches listed in 2.3.3.1 (ref. 90), 93)). Following the international discussions on statistical power and replication of chronic fish tests, the number of replicates was enhanced from two to four, resulting in the simultaneous need of two flow-through-systems for one test.

The reason for prolonging a definitive test to that duration is the ability to detect potential transgenerational or maternally transferred effects. Up to now, there is no scientific proof for the relevance of these hypothetical assumptions, but the database is very limited. Adding a further generation of test fish considerably enhances the number of tested animals. When considering animal welfare, concern of transgenerational effects and scientific plausibility, a clear recommendation is derived to change the US 2-GT test protocol. It should start with fertilized eggs of the F0 generation and end with the determination of sex ratio and first spawning of the F1 generation. This approach

- uses the same number of fish as a “conventional” FLCT
  (the US protocol needs 400 additional fish),
- includes the most critical phase of sexual development in F0 and F1
  (the US protocol includes only the sexual development of F1),
- reduces the experimentally difficult reproduction phases that add high variable data to one (the US protocol includes two),

and thus enables to test the hypothesis of maternal effect transfer more reliable and better justified. The available data is not appropriate to thoroughly identify and evaluate the significance of maternal effects for the next generation and thus not suited for the derivation of valid conclusions concerning the relevance of transgenerational effects. There is some evidence from aromatase inhibitor tests with different exposure windows that impact on egg quality of exposed parental fish reduces growth performance of unexposed filial fish. However, the effects were not as sensitive as the maternal effects themselves, and thus not relevant for regulatory decisions.

Actually, US and Japan are developing the Medaka Multigeneration Test (MMT), which is regarded superior to the fathead minnow 2-GT due to the reduced study duration and enhanced statistical power of reproduction endpoints. The different life stage phases are very similar to the proposed zebrafish phases. However, time to first reproduction is not included and the strong reduction of fish after the early life stage period weakens the power of the following endpoints considerably. IME will adapt and impove the protocol for the use of zebrafish.

2.3.4 Development of a fish full life cycle test design for pesticide higher tier risk assessment

A higher-tier zebrafish full life-cycle study design was developed to examine the potential for long-term adverse effects of pesticides on fish populations (ref. 46). To simulate more realistic exposure conditions, a static test design is used. Artificial sediment (OECD TG 219) is incorporated into the test systems to increase environmental realism but maintaining reproducibility. Three different fish life stages are concurrently exposed to a single or repeated application of the test substance according to agricultural practice. The unique study type investigates the potential of the test substance for prolonged fish toxicity, early life stage toxicity, impact on juvenile growth, impact on reproduction with special view on endocrine disruption, and early life stage toxicity of the filial generation, following peak exposure to all three introduced life stages.

Table 3: Static zebrafish FLCT with sediment: Course, estimated duration of the respective phases and important endpoints

Time after dosing	Phase (course group A)	Course group A	Endpoints*	Course group B **	Course group C **
- 4 d				Start with 50 28 d - old juveniles	Start with 30 70 d-old pre-adults
0 h	ELS toxicity according to OECD TG 210 (parental generation P)	Start with 100 fertilized eggs per egg container		Juvenile growth	
3 d		Hatch			
5 d		Feeding: breeding food			time to first
9 d		Feeding: *Artemia salina*			Spawning
2 w			Survival rate		
3 w			Survival rate		
4 w		Reduction to 50 individuals, Transfer to aquaria	Survival rate length		daily egg counts, fertilization rate
5 w	Juvenile growth	Reduction to 30 individuals	Survival rate, length	Reduction to 30	
6 w				time to first spawning	Start ELS F-
7 w				daily egg counts, fertilization rate	Generation Transfer to aquaria
8 w					
9 w					
10 w	Repro-duction	Introduction of spawning trays	time to first spawning	Start ELS F-	end P-generation
11 w				Generation	
12 w			daily egg counts, fertilization rate	end of P-generation	Juvenile growth
13 w					
14 w					
98 d	ELS toxicity according to OECD TG 210 (filial generation F)	Start with 100 fertilized eggs per vessel, transferred from the vessels of phase 2			
101 d		Hatch			
103 d		Feeding with breeding food			
107 d		Feeding with Artemia salina			
112 d			Survival rate		
119 d		Transfer to aquaria end of P-generation	Survival rate; Length, weight		end F-generation sex determination
133 d		End of FELS-study in the second exposed generation; end of the study	Survival rate Length, weight		

*only the most important endpoints listed. All apparent deviations from the controls were recorded.
** not all endpoints listed

The study is used for addressing two different concerns when focusing on fish. The first is the potential effect on populations of substances known to have a bioaccumulation potential. The second is the population effect of endocrine disruption becoming evident after peak exposure of a sensitive life stage.

Three life stages of zebrafish (*D. rerio*) (fertilized eggs, juveniles, and nearly-mature adults at the beginning of reproduction) are exposed to replicated concentrations under static conditions in large (260 L) subdivided glass aquaria with a water column of 48 cm height and a layer of 3 cm artificial sediment. Three to four untreated aquaria serve as controls.

The in-life phase is started with the simultaneous treatment of 100 fertilized eggs, 50 juveniles and 30 nearly adult fish, with each group carefully segregated in separate compartments within the test aquaria (Table 3, Figure 11). When the fish from fertilized eggs reach the age of 28 days (juveniles) their numbers are reduced to 50; similarly, when fish reaching an age of approximately 70 days (beginning of adult phase), their numbers are reduced to 30.

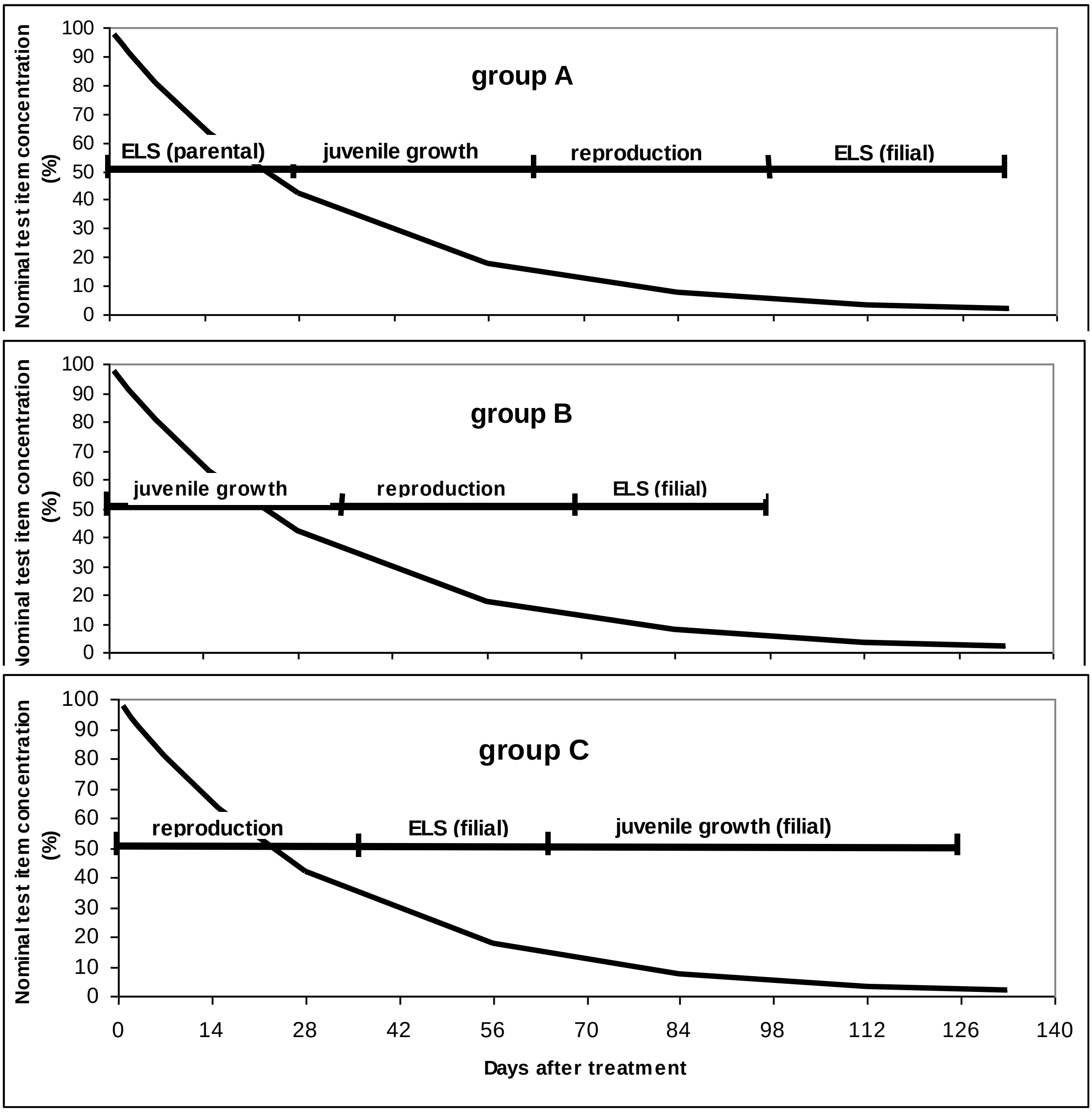


Figure 11: Exposure of the different life stages run through and performances by fish of groups A, B and C (ref. 87)

When groups are reduced, fish are digitally photographed. Survival rates and fish lengths (pseudo-specific growth) are estimated by evaluating photographs using electronically supported counting and analysis. After the last reduction to 30 individuals, glass spawning-trays are introduced and monitored daily for spawned eggs. The time until first findings of eggs is recorded. Egg production per female per day and fertilization rate is measured for at least 21 days. For the filial generation of each group, survival rates and growth until day 28 or 35 of age are observed. The filial generation of the groups exposed as nearly adults may be observed until day 70 of age, including juvenile growth and sex differentiation.

During a GLP study of this type, which is uniquely performed at the Fraunhofer IME, usually approximately 150,000 eggs are individually determined for being fertilized and approximately 9000 photographic length measurements are performed. This generates a huge data set enabling statistical evaluation with high resolution.

2.3.4.1 Bioaccumulation

Three different higher-tier full fish life cycle GLP-studies (ref. 86), 88), 89)) were conducted to examine the potential for long-term adverse effects of potentially accumulative test items on fish populations. Simultaneously to the investigation of the potential ecotoxicological effects of the test item, information about fate of the test substance in the system and about bioaccumulation under realistic conditions following various intervals of exposure was obtained. For this, an additional aquarium of the same size and test media was used, containing either all three life stages comparable to the effect study part, or only nearly mature zebra fish. This system was treated comparably to the effect aquaria with the second highest or highest test concentration of radio labeled test substance/L. Additionally, the study provided information on the behavior of the test item in the test system, i.e., an artificial water-sediment system containing fish. At the end of the in-life assessment period, the radioactivity balance was estimated.

In two studies, the test design was four replicated test item concentrations with three or four untreated aquaria serving as controls. The test systems were treated once or twice with an interval of 14 days between treatments. In the third study, the test design consisted of three test item concentrations in three replicates each, with three untreated aquaria serving as controls. The test systems were treated four times with an interval of 7 days between treatments. Three further aquaria at the highest initial target concentration were treated only twice to enable a reduction of registered applications in the case of effects at four treatments.

Exposure concentration and bioaccumulation. The initial peaks of the active ingredient deviated by less than 20% from nominal concentrations. The peak concentrations dissipated rapidly from the water column with a DT50 between 5 to 8 hours and 2 to 4 d., whereas degradation of the final percentage was sometimes slower due to specific processes, i.e., the formation of an isomer. Related to initial concentrations, a maximum bioconcentration factor of between 800 and 2000 could be anticipated under realistic worst-case exposure conditions. The number of treatments did not influence the peak concentration in water or in fish, but partly influenced the DT90 of the test substance as well as the level of residues of degradation products, measured as non-extractable radioactivity in water being twice as high after four compared to two treatments. Consequently, also the concentration of radioactivity in fish was twice as high after four compared to two treatments. Metabolism of the test item in the test systems was dominated by fish. The metabolites were more similar to those found in OECD TG 305 bioconcentration studies than to those found in water-sediment systems or mesocosm studies. Uptake and elimination trends were more pronounced in juveniles due to their rapid growth and development. Fertilized eggs exhibited considerably lower uptake rates, which resulted in lower bioaccumulation. There was no evidence that bioaccumulation was affected by test item accumulation in the sediment.

Survival and growth. At the tested concentrations, no effect of the test item on survival of any life stage of the zebrafish, neither parental nor filial, was observed. Growth was not affected in any of the treatments. The fish in all treatments met the quality criteria for control fish.

Reproduction and generation time. Within the concentration range tested, no effect on individual reproduction expressed as eggs per female and day and fertilization rate was observed.

Discussion. Throughout the entire test, fish at all life stages of all groups in all test vessels met the quality criteria for control survival and reproduction. Few exceptions for early life stage survival of the filial generation were not related to test item concentrations and could be explained by specific influence factors confirmed by repetitions of the respective study parts. Neither a retardation of reproduction nor an alteration of sex ratio was observed in any concentration tested. Thus, an effect of the test item on the parameters relevant for the population level can be excluded at the test item concentrations tested.

Conclusion. Up to the highest initial test item concentration, there was no concentration-response relationship in any of the exposure regimens. These results are considered appropriate for a higher tier risk assessment for fish, because it was consistently demonstrated that a realistic worst case exposure to the test item at initial a.i. concentrations up to the water solubility limit does not result in any effect on the life cycle of the test fish, either directly nor via bioaccumulation. This conclusion holds for peak exposure of all critical fish life stages. Thus, the NOEC is equal to the highest concentration tested. Accordingly, no adverse effects on fish populations are expected to occur up to and including the highest initial test item concentration.

2.3.4.2 Endocrine disruption: Aromatase inhibition

Three studies were performed with fungicides known to inhibit aromatase activity in mammals for assessing the risk of endocrine disruption in fish at a minimum of uncertainty (ref. 87), 91), 92)). Exemplarily, one study is presented (ref. 87). Five concentrations were tested in two replicates with four untreated control aquaria. Beside the generally investigated endpoints, concentrations of the yolk protein component VTG were measured in blood samples of adult male and female fish. The gonads of the same fish were investigated histologically for abnormal appearance by Helmut Segner, University of Bern, Switzerland.

Survival and growth. At the tested treatment levels, no effect on survival of any life stage of zebrafish, parental or filial, could be observed. Growth was retarded in the early life stages at the highest treatment level, whereas juvenile growth was not affected in any of the treatments.

Reproduction. Except the highest treatment level, no effect on individual reproduction expressed as eggs per female and day and fertilization rate could be observed. The effect at the highest treatment level was caused by the retardation of spawning.

Sexual development and manifestation. The respective endpoints were measured at the end of the parental generation. Sex ratio and VTG level were the most sensitive endpoints, whereas gonad histology was less sensitive. Sex ratio was significantly shifted to males, especially in fish exposed during sexual development (started as fertilized eggs or juveniles, clear effects at the two highest concentrations). VTG concentrations in blood plasma were significantly reduced in females, if sufficiently present, with the most sensitive NOEC of the third concentration in fish exposed as (pre-) adults. In males exposed during sexual development, the NOEC concerning VTG was as low as the fourth concentration, but not regarded as relevant (see discussion).

Discussion. The specific mode of action of the test substance is known to also inhibit the aromatase of vertebrates, the enzyme being responsible for the formation of 17β-estradiol out of testosterone and thus influences the manifestation of the sex, the development of gonads and specific sex performances like VTG production, egg and sperm production, and mating behavior. Because of its specific development of the male gonads from protogyn gonads, zebrafish can be regarded as particularly sensitive to aromatase inhibition: surpassing a threshold ratio of testosterone / 17β-estradiol seems to be the cue for transforming protogyn to male gonads.

The results of the study consistently demonstrate that there are two different patterns of effects depending on the time window of exposure in relation to sexual development:

1) When the complete sexual development occurred during exposure, high concentrations of the test substance led to sex reversal of sensitive females to phenotypic and functional

males. The NOEC for alteration of the sex ratio was found to be the third initial concentration. Fish exposed as juveniles to the treatment shortly after the start of sexual development reacted most sensitively to this parameter. Fish introduced as fertilized eggs first had to run through the early life stages before main sexual development started; by this time concentrations of the test substance had decreased by about a factor of two. Accordingly, the ECx values for the parameter sex ratio were about a factor of two lower (based on initial concentrations). For the reduction of female VTG concentration, the NOEC could not be determined for fish introduced as juveniles, since the only remaining female exhibited VTG concentrations similar to controls. For fish introduced as eggs, the NOEC was determined to be the second concentration. The endpoint “VTG concentration in males” appeared to be a very sensitive endpoint (NOEC: fourth concentration), probably because of its irrelevance for populations. There is no biological need for compensatory regulation of 17β-estradiol and VTG.

2) Fish introduced as adult and partly pre-adult fish performed sexual development mainly, but not totally before exposure, demonstrated by the start of spawning 11 days after treatment. Only the highest concentration led to sex reversal of some (lately developing) females to males (NOEC: second concentration). This reversal was not always completely phenotypic or functional. Although not statistically significant there were hints to gonad abnormalities and decreased fertilization rates in one replicate each at the highest three concentrations, which seemed to be correlated. At high test substance concentrations, the high number of females left resulted in a more powerful statistics on VTG concentrations, being the most sensitive endpoint in this group with a NOEC at the third concentration. The filial generation of the fish introduced as (pre-) adults exhibited a clear effect on fish length at the highest treatment level (NOEC: second concentration) which cannot be consistently explained as a direct effect of the test item on the fish larvae directly. Instead, it is concluded that the reason for the very early deficiency could be a decreased yolk supply of eggs caused by reduced VTG levels in the female parental blood plasma. However, the growth retardation could be completely compensated during the juvenile growth period.

If the effects at the different life stages are compared by time-weighted averages for the exposure of the most sensitive stages (concerning the most sensitive effects), then similar effect concentrations could be calculated, supporting the conclusions that the study results are highly reproducible. The EC50 for the alteration of sex ratio based on time-weighted averages of concentrations during sexual development was calculated to be very close in fish exposed as eggs and in fish exposed as juveniles between the second and third initial test concentration. The EC50 for the reduction of VTG in females based on time-weighted averages of concentrations from pre-adult stage to termination was calculated to be very close in fish exposed as eggs and in fish exposed as pre-adults, slightly higher than the second initial test concentration. All four EC10 values (calculated from the time-weighted average concentration during the most susceptible life period) were around the fourth initial test concentration.

The number of fertilized eggs per female can serve as an integrative figure for population effects caused by effects on generation time, reproduction and sex ratio. A reduction below the control range of the reproduction rate could be observed at the two highest concentrations, being consistent with the NOECs concerning sex ratio and VTG reduction of the third concentration.

Conclusions. The study results are suited for a higher tier risk assessment for fish, since they are consistent for interpreting the mode of action of the test substance with respect to different effect patterns depending on the time window following a peak exposure. It is possible to derive a No Observed Ecologically Adverse Effect Concentration (NOEAEC) based on initial concentrations. The potential of the test method was demonstrated by two further studies (ref. 91), 92)).

The most sensitive endpoint determined, effects on male VTG concentrations with a NOEC of the fourth concentration, is not regarded to be of relevance for fish populations; the respective NOEC for female VTG, which would be of relevance for the offspring, is the second test concentration. The next most sensitive endpoint is an effect on the sex ratio with a NOEC at the third test concentration. On the experimental population level, that NOEC is confirmed by the cumulative numbers of fertilized eggs per group.

Thus, in consequence, the NOEC relevant with respect to adverse effects (NOEAEC) is the third test concentration. This is confirmed by the calculated EC10 values. When focusing on fish performing their sexual development under exposure, the lowest EC10 for sex reversal was calculated to be slightly below the third test concentration and comparable to the lowest EC10 for reduced VTG levels in females exposed as sexually mature fish. However, a slight sex reversal of less than 20% most probably does not represent an adverse effect on fish populations and the effect of VTG reduction on egg quality only became manifest at the highest test concentration (above the EC50 of VTG reduction). Even that effect (measured as growth retardation of the filial early life stages) could be compensated by a natural food supply. Accordingly, no adverse effects on fish populations – including potential endocrine activities – can be expected at initial maximum concentrations of the test substance as high as the third test concentration or less (= NOEAEC).

Scientific perspective: For demonstrating consistency of data and certainty of results and for showing scientific correlations rather than regulatory thresholds, the following evaluation approach was chosen. Sensitivities of fish exposed at different life cycle periods were compared by relating effect data to time-weighted averages of concentrations measured during exposure of the life stages most susceptible towards the most sensitive effects (sex ratio and VTG level in females) (Figure 12). By doing that, the different effect concentrations were normalized for the real exposure dose of the most susceptible life stage. EC50s for the two endpoints were calculated separately for each relevant exposure group. The EC50s for the alteration of sex ratio in fish exposed as eggs and in fish exposed as juveniles, related to the time-weighted averages of concentrations measured during sexual development, i.e. between day 20 of life and the start of spawning, differed by less than 10%. The same consistency was achieved for the reduction of VTG in females in fish exposed as eggs and in fish exposed as pre-adults, related to the time-weighted averages of concentrations measured during reproduction, i.e. between ten days before the start of spawning and blood sampling. For the sex ratio as well as for VTG reduction in females, the EC20 as a threshold concentration for effects was similar for both endpoints, supporting the following conclusions:

- The assumptions of the most sensitive periods for sexual development and VTG production were confirmed.
- The study results were shown to be highly reproducible.
- The concentration-response-relationship for the alteration of the sex ratio caused by exposure during sexual development was clearly steeper than that for the reduction of female VTG caused by exposure as pre-adult and adult fish, indicating to a threshold in developmental regulation (e.g. ratio of testosterone / 17β-estradiol).
- The non-physiological alteration of the VTG concentration was demonstrated to be a precise biomarker for the threshold of effects relevant for the population level (alteration of sex ratio).

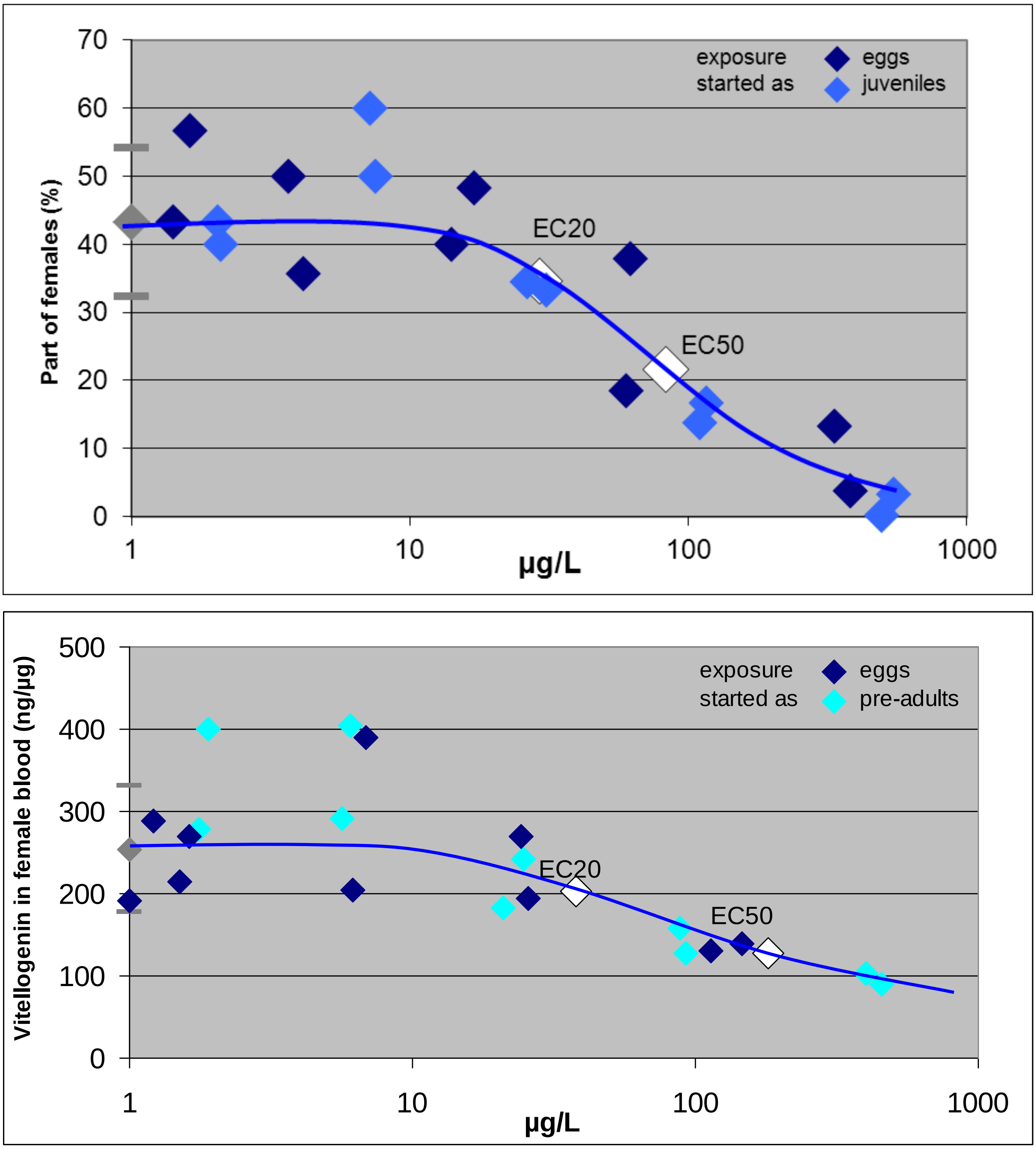


Figure 12: Effects of an aromatase inhibitor on sexual development of zebrafish
Concentrations given as time-weighted averages of measured concentrations during sexual development (day 20 of life until start of spawning)
Upper: Proportion of females after completed maturation
(fish starting as pre-mature adults already had left the most sensitive stage).
Lower: VTG concentration in female blood plasma, expressed as proportion of females (fish starting as juveniles nearly totally developed as males, see upper).

Figure 13: Test systems for zebrafish life cycle tests
Left: flow-through system for continuous exposure and investigation of intrinsic toxicity. Right: static system including sediment for simultaneous peak exposure to three different life stages, used in higher tier risk assessment for pesticides.

2.3.5 Quality criteria of life cycle test with zebrafish

Within the OECD discussions of a fish full life cycle or two generation test (OECD, 2003), zebrafish is one of three species in the focus. Due to the dominance of American needs in the last decades, early life stage and juvenile growth tests have been performed mainly with rainbow trout and fathead minnow. Thus, information is scarce on data variability concerning the endpoints of a zebrafish two-generation study. During the last five years, we conducted several zebrafish full life cycle studies with different objectives and study designs. The anonymized non-effect data were used for a statistical evaluation of data variability (ref. 56). The comparison focuses on individual replicates of three main test types of zebrafish FLCTs: Flow-through tests (Figure 12, left), semi-static tests with three renewals a week, and static tests including sediment (Figure 12, right), all starting with 100 fertilized eggs per replicate. The data set includes all control replicates and the treated replicates without test item effect. When identified study conditions were responsible for enhanced data variability, data were excluded.

2.3.5.1 Results

Early life stage total success. Means and variances of flow-through and static tests were comparable. The mean of 80% survival (SD = 8.8%) represents approximately 19000 individuals. The quality criterion of the OECD TG 210 of 70% mean control post-hatch success was met by all 23 data sets except one. 8% of replicates were below 70% total success. The semi-static tests (all non-GLP) showed clearly worse results with a lower mean and higher variability.

Reproduction. For total eggs per female and day, means and variances of all test types were comparable (205 replicates with approximately 15 females and 20 daily egg counts ≈ 2.2 million eggs). The mean of 37 eggs per female and day is within a high range of variability. For 160 replicates (78%), eggs were counted to be between 20 and 50. As quality criterion a mean control egg number per female and day of 20 can be derived (6% of the replicates were lower, Figure 14). For fertilization rates, the means and variances of flow-through and static tests were comparable. Rates were high (mean of 93 %) with very low variability (relative standard deviation below 5%; approximately 2 million eggs, each individually checked for being fertilized shortly after spawning). As quality criterion a fertilization rate of 85% can be derived (2.5% of replicates were lower). The semi-static tests (non-GLP) showed clearly worse results with a lower mean and higher variability.

Sex ratio. For the determination of sex ratios, only fully developed gonads were used. The means and variances of flow-through, static and semi-static tests were comparable (~4000 fish checked for sex by gonad inspection after observation of reproduction). The portion of females was ~50% with high variability. Less than 5 % of the replicates were outside a range of 25-75%. Normal reproduction was performed at a range of 35-65% females.

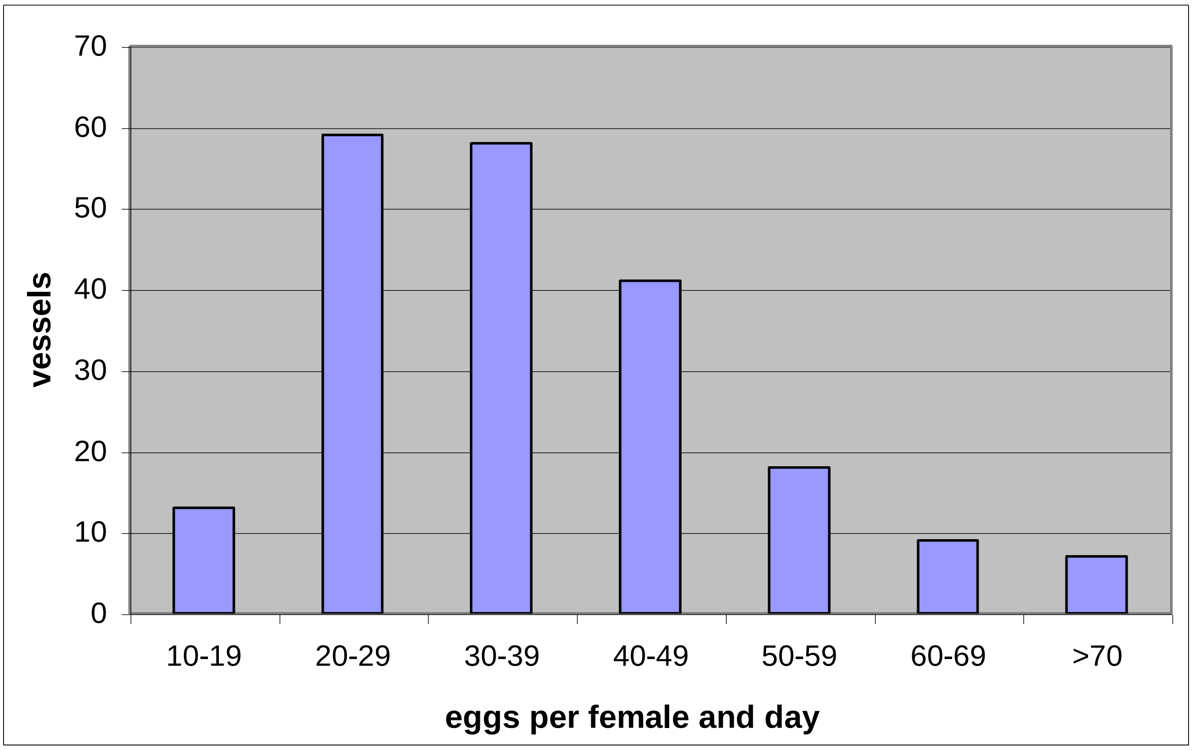


Figure 14: Number of eggs per female and day produced in zebrafish life cycle tests
Data from controls and unaffected test concentrations of tests performed at the Fraunhofer IME; means of 20 daily counts per vessel, divided by the number of females

2.3.5.2 Conclusions

The comparison could demonstrate the endpoint variability of zebrafish life cycle tests for the laboratory-bred strain used. For early life stage total success, the quality criterion of the OECD TG 210 is feasible. By using two replicates per concentration and four controls, a 25% deviation from controls is probably statistically significant. For eggs per female per day, a quality criterion of 20 is suggested. Due to the high variability, at least a 75% deviation from controls is needed to be statistically significant. For fertilization rate, a quality criterion of 85% is suggested. Due to the low variability, a 10% deviation from controls is probably statistically significant. For sex ratio, a normal range of 35-65% is suggested. A deviation of more than 50% from control means is probably statistically significant.

2.3.6 Different endocrine modes of action and endpoint sensitivity

The first years of research on endocrine disrupting chemicals were strongly focused on estrogenic effects. Estrogen receptor assays were developed to identify the estrogen receptor (ER) agonists responsible for feminization. For fish, FLCTs were conducted with chemicals of concern (mainly contraceptive pharmaceuticals, groups of industrial chemicals and phytoestrogens) to identify the hazard potential for fish populations (see 2.3.2) and improved with respect to endocrine endpoints (see 2.3.3, 2.3.5). With growing database about estrogenic effects, the interest in other endocrine modes of action increased.

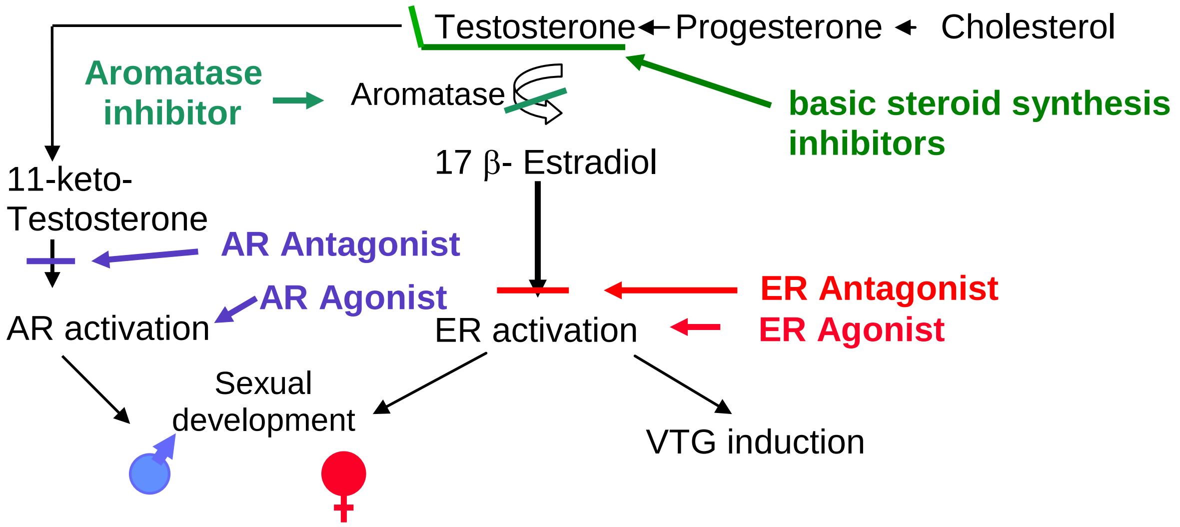


Figure 15: Scheme of direct Modes of Action of endocrine disruption, strongly simplified

2.3.6.1 Aromatase inhibition of DMI fungicides

Especially masculinization by aromatase inhibition (Figure 12) became an issue due to the intended MoA of the world's most important fungicide class. DMI fungicides (mostly azoles) inhibit the ergosterol synthesis in the fungi cells, the target enzyme being partly analogous to the aromatase and demethylases in the testosterone synthesis pathway of vertebrates and other animal classes with sexual steroids. Consequently, for aquatic risk assessment of DMI fungicides, the agrochemical industry was forced to conduct FLCTs with fish (see 2.3.4). As these studies are time-consuming, the registration of the fungicide use has to be postponed for one year.

In a comparative desk study on confidential data of all relevant agrochemical companies (ref. 60), the most sensitive endpoints of FLCTs with zebrafish, fathead minnow and sheepshead minnow were compared with the most sensitive endpoints of long-term and chronic standard tests with rainbow trout, fathead minnow or sheepshead minnow. For all substances, aromatase inhibition seems to be the main MoA. The sensitivity of the population-relevant endpoints sex ratio, growth and reproduction was mostly close. All three endpoints deviated by one concentration step at maximum in six of nine valid tests. The higher deviations in the other three tests could be explained by duration of exposure or irregular sex ratio in the controls. The main weakness of the FLCT data set is the insufficient determination of the sex ratio in many studies. Effects on reproduction sometimes may be due to a shifted ratio of functional sex. It is obvious that juvenile growth during the period of sexual maturation is related to the shift of sex ratio and probably due to enhanced need of energy for compensatory reactions, e.g. up-regulation of aromatase. For one substance, two species were tested. The sensitivity was equal, but the most sensitive endpoint in fathead minnow was growth whereas in zebrafish it was sex ratio. It is assumed that fathead minnow is more potent in compensatory aromatase regulation.

When comparing the most sensitive FLCT NOECs of the seven substances with the most sensitive growth NOECs from long-term and chronic standard tests, in six cases the standard NOEC is as sensitive as or more sensitive than the FLCT NOEC (Figure 16). The only substance deviating from this is the substance with the lowest acute-chronic ratio, with other aquatic species being more relevant for the risk assessment. There are two potential reasons, which can be discussed. When more than one standard study was available, only the most sensitive endpoint was included. Thus, the probability of lower findings in standard tests increased. The most common species for pesticide ELS and JG tests is rainbow trout. The long test duration in an ELS (around 90 d) and the high absolute growth are reasons for a relatively high sensitivity. Thus, when accounting for these aspects and including a statistical evaluation of the ratio of all standard test NOECs : most sensitive FLCT NOEC, an extrapolation factor of 5 to a FLCT NOEC seems to be justified for a preliminary risk assessment based on standard tests. This preliminary factor has to be confirmed or reduced by further studies that have to be performed in parallel to registration. The upper 99.9 percentile of the distribution of NOEC ratios for all seven substances (including substance 14, Figure 16) was 4.6.

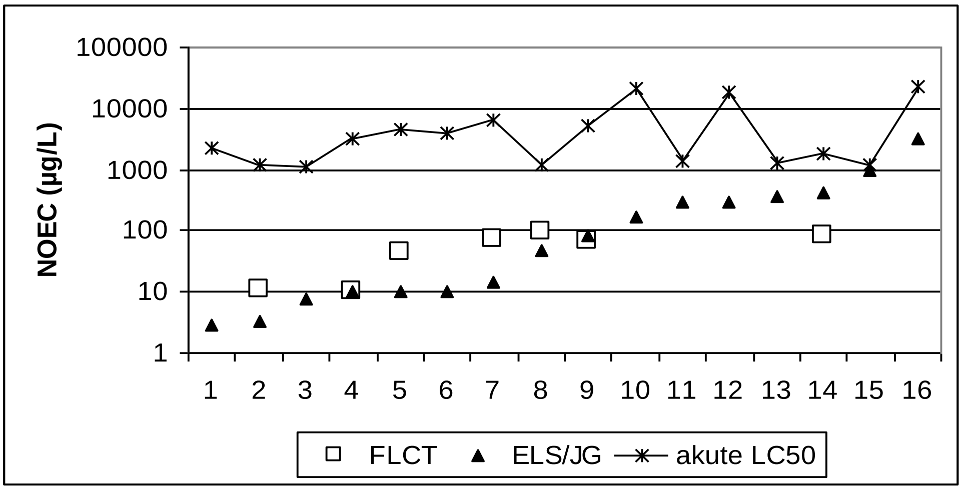


Figure 16: Most sensitive NOEC data from early life stage (ELS) and juvenile growth (JG) tests compared with the most sensitive endpoints of valid full life cycle tests (FLCT) and acute LC50s. Order of substances (1-16) according to ELS/JG sensitivity

2.3.6.2 Further sexual endocrine MoAs

Whereas the data situation for ER agonists and steroid synthesis inhibitors (especially aromatase inhibitors) is acceptable for generalizing assessments, FLCT data on AR interactions and ER antagonists are missing or scarce. In order to fill the gaps by conducting appropriate state-of–the-art studies with the same fish strain used for the other MoAs, an UBA project was initiated. In cooperation with the University of Heidelberg, 2-GT tests with IME zebrafish[3] were performed using trenbolone as AR agonist (Heidelberg), the pesticide flutamide as AR antagonist (IME) and the pharmaceutical tamoxifen citrate as ER antagonist (IME). All histological analyses were performed at the University of Heidelberg, all blood samplings and analyses were performed by the IME staff. At the same time, short-term assays following the protocol for the fish screening assay (now 21d-FA, OECD 2009) were performed for all substances at the IME for comparison of biomarker responses as indicative endpoints (Figure 17) after short-term and long-term exposure. The report of the recently finalized studies is in preparation (ref. 62).

---

[3] *Danio rerio* (Hamilton-Buchanan 1822) (Teleostei, Cyprinidae), laboratory inbred since > 15 years. Origin of the used strain of zebra fish: West Aquarium GmbH 37431 Bad Lauterberg, Germany

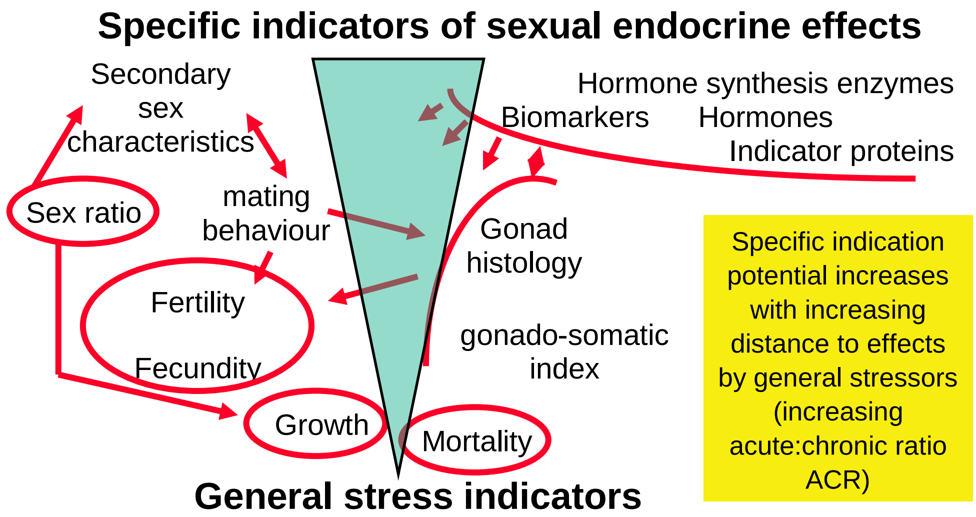


Figure 17: Interrelation between endocrine population-relevant (encircled) and indicative endpoints

2.3.6.3 Desk study on sexual endocrine effects in fish full life cycle tests

In addition, a literature study sponsored by the UBA was performed in cooperation with ECT, Flörsheim, Germany, with the scope of the reduction of uncertainty in assessing fish FLCTs by a systematic collection and evaluation of existing data (ref. 61). Population relevant and indicative endpoints specific for basic sexual-endocrine modes of action (MoA) should be identified and compared with regard to sensitivity. For this, the direct sexual-endocrine effects were again pragmatically classified as interactions with the estrogen or androgen receptor (agonism or antagonism) or as inhibitions of the steroid synthesis (basic steroid synthesis or aromatase inhibition). The database consisted of zebrafish full life cycle and 2-GT tests by the Fraunhofer IME and from literature, mainly on fathead minnow (*Pimephales promelas*), Japanese medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*). Based on the results of the evaluation it was tried to classify FLCT data of the UBA regulatory database with unknown MoA according to the sensitive endpoints reported with respect to potentially sexual endocrine effects.

From the data set, some statements can be derived regarding the sensitivity of different population relevant and indicative endpoints. However, it has to be underlined that an adequate database concerning tested substances, fish species and applied test methods is only available for estrogen receptor agonists and aromatase inhibitors, especially azole fungicides.

Comparison of species. The sexual development of fish species exhibits huge variation in determinants (genetic sex versus environmental cues), timing of differentiation (embryonic determination up to multiple changes of sex after maturity) and succession (development direct, via indifferent stages or by development of protogyn gonads). This potentially results in differences in the susceptibility to endocrine disruption regarding the sensitive phase of exposure, the potency of compensatory regulation and the manifestation of effects. The investigated fish species seem to be principally comparably sensitive towards hormone receptor interactions (Table 4). The manifestation of effects may differ, i.e., in Medaka and Fathead Minnow, ER agonists cause feminization, whereas Zebrafish are arrested in their male protogyn development phase. Zebrafish seem to be more sensitive to masculinization by aromatase inhibition and probably ER antagonists. Fathead minnow seem to be more potent in compensatory aromatase regulation, but also to be more sensitive in a reduction of juvenile growth (higher absolute growth). The cyprinids zebrafish and fathead minnow seem to be determined by genetic markers which are not located on a sex chromosome, but on heterosomes. Thus, medaka, which has sex chromosomes, can be individually evaluated for sex reversal, whereas in the other two species the evaluation has to be based on statistical distributions around the mean sex ratio and consequently is less precise. According to Les Touart, USEPA (personal communication), the identification of the heterosomal sex markers in fathead minnow (and zebrafish) is a matter of time. In the USA, fathead minnow is the preferred species for short-term assay, whereas medaka is focussed on in two-generation tests, as fathead minnow tests need considerably more time and efforts. When looking at wide ranges of toxicological action, Medaka often is less sensible than fathead minnow or zebrafish, which might be a result of a higher metabolic potential as described for guppy (*Poecilia reticulata*) (ref. 37).

Table 4: Comparison of inter-species sensitivity to an ER agonist (ethinylestradiol)

Endpoint		**Effects of ethinylestradiol** Species	LOEC	
VTG in males (short term tests)		Zebrafish	1 ng/L	
		Fathead minnow[1]	1 ng/L	
		Rainbow trout[2]	0.3-1 ng/L	
Most sensitive population-relevant FLCT-effects:	Fertility		EC10	EC50
		Zebrafish	0.6 ng/L	1 ng/L
		Fathead minnow[3]	0.3 ng/L	1 ng/L
	Sex ratio		NOEC	EC50
		Fathead minnow[4]	1 ng/L	4 ng/L
		Medaka[5]	1 ng/L	10 ng/L

[1]Pawlowski et al. (2004), [2]Sheahan et al. (1994), [3]Parrot et al. (2005), [4]Länge et al. (2001), [5]Scholz et al. (2004)

Sensitive phase for exposure and effects. The most sensitive exposure period for interactions with the estrogen receptor (ER) is the sexual development phase. The most sensitive manifestation endpoint of **ER agonistic** effects is the reduction of the fertilization rate, the sensitivity of juvenile growth and sex manifestation being very close. The VTG measurement is comparably sensitive. **ER antagonists** seem to enhance the androgen level by feedback stimulation and are able to induce male development in zebrafish.

For Androgen receptor (AR) interactions, VTG is not suited as a powerful biomarker. For the **AR antagonist**, the most sensitive population relevant endpoint is fecundity (spawned eggs per female). The most sensitive indicative endpoint is the enhanced sexual steroid 11-kT. The most sensitive exposure period cannot be derived from a flow-through test. However, as the effect became manifest in the reduced number of spawned eggs, most probably caused by the reduction of male mating behavior and being reversible when the concentrations dropped, this can be regarded as evidence for the reproduction phase being the most sensitive exposure period. Regarding the **AR agonists**, sexual development is the most sensitive exposure period, the sex ratio being the most sensitive endpoint. Sufficiently sensitive biomarkers were not identified. **Aromatase inhibition** causes most sensitive effects when occurring during the sexual development phase. A sexual endocrine effect becomes manifest in different endpoints, the sensitivity being as close as factor 3 in NOECs in all six valid and comparable tests: a shift in the sex ratio, a decrease of fecundity and a reduction of juvenile growth. If measured, VTG reduction in females as indicative endpoint was comparably sensitive.

Conclusion. Consequently, the identification of a concern of sexual endocrine effects should include hints of the most probable MoA to direct the test designs and focus the test strategy. This entrance into the test strategy was not part of the study but will be discussed in a working group formed by representatives of industry, UBA and Fraunhofer. With respect to the test species, each has its advantages in disadvantages in endocrine testing (Table 5).

Table 5: Advantages (+) and disadvantages (-) of the life cycle test species. 0 = average or data insufficient.

Endpoints	**Fathead minnow**	**Medaka**	**Zebrafish**
Statistical sensitivity			
Early life stages:			
Hatch	**+**	**-** (time; movement)	**+**
Survival	**+**	**+**	**-** (early feeding)
Juvenile growth	+ (absolute growth)	0	0
Reproduction (variability of fecundity)	- (male dominance)	0	**+** (group spawning; absolute numbers)
Secondary sex characteristics	**+**	**+**	**-**
Genetic sex	**-**	**+**	**-**
Blood sampling	**+**	**-**	0 (cardiac puncture)
Egg collection	**0**	**-**	**+**
Physiological sensitivity			
Generally		**-** (less sensitive) (metabolism?)	
ER or AR interactions	**+**	**+**	**+**
Aromatase inhibition	**+** juvenile growth	0	**+** sex ratio
Replication	**-** (due to reproduction)	0	**+**
Test duration	**-**	**+**	**+**
The evaluation is subjective and can partly be compensated by the skills of the performing laboratory.			

2.3.7 Can short-term tests indicate endocrine disruption with sufficient sensitivity?

A main research issue at Fraunhofer IME during the last years was the comparative assessment of population-relevant and indicative endpoints sensitive to direct modes of action (MoA) of sexual endocrine disruption. Selected substances representative for all relevant MoA were investigated in short-term assays (duration: 14 to 21 days; endpoints: biomarkers VTG and 11-kT in the blood plasma, fecundity and fertility) and full life cycle or 2-GTs (ref. 21), 49), 58), 61)). At the same time, the Fraunhofer IME was involved in the development and validation of the OECD 21d fish assay (21d-FA, OECD 2009) and acted as lead laboratory zebrafish for the phase 1b of the validation exercise.

The objective of this work was the identification of effect syndromes specific for the endocrine MoA to be able to focus the test strategy and endpoint observation. At the same time, it should be clarified whether biomarkers as indicative endpoints in short term assays like the 21d-FA are always sensitive enough to exclude false negative responses. This is crucial for the use of the 21d-FA in a tiered testing approach, as a negative response results in stopping further tests. In this regard, and in contrast to Hutchinson et al. (2006), the endpoints of a 21d-FA are used as traffic lights and not as signposts.

Effect syndromes. Only for the ER receptor agonists, the effect syndrome is clear and indicative because it is dominated by the direct interaction with the receptor: A sexual endocrine effect most sensitively and specifically becomes manifest in the reproductive parameter fertilization rate (Table 6) (ref. 27)). However, the sensitivity of juvenile growth and sexual development (time to first spawn) are close to fertility. As indicative endpoint, the blood plasma concentration of the yolk precursor protein VTG provides as sensitive results as the most sensitive population relevant endpoint, which can easily be shown, as the induction of VTG in both sexes can powerfully be measured, particularly in exposed males. This syndrome was shown for strong and for weak estrogens.

Table 6: Effect syndromes in zebrafish after exposure to various sexual endocrine acting substances

Mode of Action	Chemical (group)	Short-term test			Life cycle test		
		VTG	**11-kT**	**Reproduction**	**VTG**	**11-kT**	**Population**
Strong	EE2	+	–	Fertility	+	n.a.	Fertility, time to 1st spawning
ER agonist	Alkylphenol	+	–		+	n.a.	
Weak	BPA	+	–		+	n.a.	Fertility
ER antagonist	Tamoxifen	–	+		–	+	Fecundity, sex ratio
AR agonist	Trenbolone	0	0		0	0	Sex ratio
AR antagonist	Flutamide	0	+	Fecundity	–	+	Fecundity
Aromatase inhibitor	Fadrozole	–	+	Fecundity	n.a.	n.a.	n.a.
	Azole fungicide	–	(–)	Fecundity	–	(–)	Sex ratio
Testosterone synthesis inhibitor	3,4-DCA	–	–		n.a.	–*	Fecundity (ELS)
	Atrazine	–	(–)		–	n.a.	Fecundity (ref. 57)

VTG: vitellogenin; 11-kT: 11-keto-testosterone; ER: estrogen receptor; AR: androgen receptor; EE2: ethinylestradiol; BPA: bisphenol A; 3.4-DCA: 3,4-dichloroaniline; n.a.: not analyzed
Alkylphenol: pentylphenol in short-term test, p-tert-octylphenol in life cycle test
*Allner (1997)

For other MoAs, effects are less indicative, because indicators respond less pronounced and direct effects interfere with feedback inhibition or stimulation. It is more difficult to detect a significant VTG reduction caused by the ER antagonistic MoA, which is mainly possible in females, as VTG levels in males are very low anyway. For reproducibly characterizing ER antagonism effects, the database has to be extended. For all MoAs except the AR agonist (trenbolone), the most sensitive biomarker responses (VTG or 11-kT) were as sensitive as the most sensitive population-relevant endpoints. Despite the efforts of harmonization of the generation, evaluation and interpretation of histopathological findings there are still uncertainties. The matching of sensitivity with that of population-relevant endpoints is less clear than for the biomarkers. However, histopathology is a valuable support for the interpretation of results, especially when additional modes of action are of importance (e.g. liver toxicity) or biomarkers are not sufficiently sensitive (AR agonists). For the endocrine effects caused by the least specific substances 3,4-DCA and atrazine, the acute chronic ratio is below 20 (see chapter 0) and other effects like early life stage survival are more sensitive, indicating the influence of general toxicity that of cause also includes the endocrine system.

When comparing the sensitivity of biomarker responses from short-term assays like the 21d-FA with the effect data from population-relevant endpoints of full life cycle and two-generation test with zebrafish, the VTG response in short-term tests was as sensitive as the most sensitive population-relevant FLCT / 2-GT endpoint except for the AR interactions (Table 7). For the AR-anatgonist, enhanced 11-kT, which can be measured using the same blood sample as used for VTG measurements, indicates feedback stimulation, which is as sensitive as mating and fecundity as most sensitive population-relevant endpoints.

Table 7: Effect concentrations in zebrafish following exposure to different sexual endocrine acting substances. Short-term test: Most sensitive biomarker; life cycle test: most sensitive population-relevant endpoint

Mode of action	**Chemical (group)**	**Short-term test**		**Life cycle test**	
		LOEC	**EC50**	**LOEC/EC10**	**EC50**
ER agonist **strong**	EE2	1.1 ng/L	3.1 ng/L	0.6 ng/L	1.1 ng/L
ER agonist **weak**	BPA	375 µg/L	639 µg/L	390 µg/L	1410 µg/L
ER antagonist	Tamoxifen*	0.1 µg/L	3.0 µg/L	1.2 µg/L	8.6 µg/L
AR agonist	Trenbolone		**> 90 ng/L**	**3 ng/L**	
AR antagonist	Flutamide	47 µg/L	700 µg/L	256 µg/L	740 µg/L
Aromatase inhibitor	Azole fungicide	22 µg/L	49 µg/L	19 µg/L	86 µg/L
Testosterone synthesis inhibitor	3,4-DCA	96 µg/L	152 µg/L	117 µg/L**	141 µg/L**
	Atrazine (ref. 57)	67 µg/L	730 µg/L	n.d.	30-4270 µg/L

n.d. = not determined
* as citrate
** Ensenbach (1991), fecundity (most sensitive endpoint: early life stage survival; different MoA)

2.3.8 Ongoing and future work: A testing strategy for the assessment of endocrine disruption

Based on the presented results a proposal can be derived for a tiered test strategy for the aquatic hazard assessment of potentially sexual endocrine disrupting chemicals (Table 8). The following criteria have to be considered:

- The possibility of false negative findings should be minimized as far as possible, as no risky endocrine acting substance should be missed in the observation and assessment scheme. This also includes the option that all substances should be screened for endocrine potential and not only substances with established concern.
- On the lower tiers, efforts should be limited to the generation of data that are necessary for triggering or stopping higher tier tests. Only limitation permits testing of a wide range of substances. It is not the scope of screening tests to satisfy scientific research on MoAs.
- The possibility of false positive findings should be reduced. However, the minimization of false negatives has priority.
- The primary strategy by EDSTAC to implement three *in vivo* tiers (fish screening assay, short-term reproduction/partial life cycle test and 2-GT) is not appropriate regarding the criteria mentioned above. Either the tier 2 test (short-term reproduction/partial life cycle test) covers the most sensitive exposure and effect manifestation phases and can replace the definitive test, or not. In the latter case, it will lead to a false negative response. When regarding the predictive potential of the biomarker responses, there is no need for an intermediate tier anyway.

Table 8: Tiered test strategy for the aquatic hazard assessment of potentially sexual endocrine disrupting chemicals (ref. 49)

Tier	**Methods**	**Objective**
Basic tier	SAR; Information from toxicological or effect studies; ACR; *in vitro*-Screening (e.g. receptor assays, enzyme inhibition assays)	Identification of potential of sexual endocrine disruption
Tier 1	*In vivo*-Screening (OECD 21d-FA) including biomarkers VTG und 11-kT; for hints to AR agonists (basic tier): marker to be identified. Appropriate fish species	Identification of endocrine disruption at relevant concentrations in water
Tier 2	2-GT starting with fertilized eggs of F0, ending with sexually mature F1	Hazard assessment (relevance for fish populations)

IVA/VCI/UBA workshop. In December 2007, the workshop “characterization of endocrine effects in fish” on behalf of the agrochemical producers association IVA, the association of the German chemical industry and the UBA took place in Berlin. Delegates of the industry and authorities and invited experts from research institutes (including the IME) discussed seven issues with the aim to agree upon a test strategy appropriate for urgent regulatory needs as well as for feeding the ongoing work of the OECD.

Issue 1: What are the reasons for a concern of endocrine disruption of environmentally relevant substances (e.g. pesticides, industrial chemicals)?

Issue 2: Weight of evidence: What are the rules for a decision to conduct a 21d-FA?

Issue 3: What is the objective of a 21d-FA: Clarification of the MoA? Use of the results in regulatory decisions?

Issue 4: Which population-relevant endpoints should be covered in a definitive test?

Issue 5: Are indicative endpoints to be measured in a definitive test?

Issue 6: Which test is more appropriate: a FLCT or a 2-GT?

Issue 7: Can abbreviated FLCTs be appropriate? Are conclusions from indicative endpoints sufficiently safe to simplify a FLCT or 2-GT? What are the pre-requisites for a replacement of the definitive test (FLCT or 2-GT) by a fish sexual development test (FSDT) or a partial life cycle test including reproduction (PLCT)?

Basic tier. The issues 1 and 2 were combined to agree upon main criteria by which a concern of sexual endocrine effects of a substance can be raised and a tier 1 test be triggered.

- If the acute : chronic ratio (ACR) in vertebrates (fish, birds, mammals) for growth as chronic endpoint is high, and if no other MoA than a sexual endocrine one is obvious, there is reasonable concern of sexual endocrine disruption. From experience with usual ACRs by industry and research, an ACR>20 is proposed as threshold value.
- Proposal for appropriate criteria to evaluate *in vitro* test results: Comparison of the sensitivity relative to a positive control; distance to general toxicity.
- Screening of tests on birds and mammals for hints of sexual endocrine disruption
- Read across
- (Q)SAR
- Further findings in literature or biomonitoring

Based on the data collected for all criteria, an integrated assessment should lead to the decision whether a 21d-FA is to be performed or not. This can be a case-by-case decision in dialogue between notifier and authority. The criteria need further elaboration, possibly addition, ranking and trigger for the performance of a 21d-FA. For this purpose, a working group will be established that includes toxicological expertise and consists of delegates from industry, authority (UBA) and research (Fraunhofer IME and ITEM)[4]. The objective is to prepare a checklist for confirming a concern of sexual endocrine effectiveness of a substance. Information as mentioned by the REACH TDG (appendix 5) and by the IVA paper „Some IVA thoughts on fish endocrine testing“ (2007) will be considered by the working group. The group will recommend whether a validation of the checklist by a research project is necessary.

Tier 1 (issue 3). The 21d-FA is able to identify potential endocrine disruptors intrinsically, which cannot be managed by juvenile growth or early life stage tests due to the lack of appropriate endpoints or sensitive phases of exposure. The 21d-FA is not able to quantify population-relevant effects. However, a negative response should enable stopping further tests, if all sexual endocrine MoAs are addressed either by an extended 21d-FA (+ 11-kT, histology) or by appropriate *in vitro* assays. The short-term reproduction assay (FSTRA) is supported by the USA instead of the 21d-FA. This might be an option to include all

[4] In a project sponsored by the UBA (FKZ 206 67 448/05 “Development of strukture and risk based methods for the identification of chemicals with endocrine effect potential for the priorisation in the REACH notification process”), Fraunhofer IME (ecotoxicology and QSAR, in cooperation with M. Nendza) and ITEM (toxicology and experimental medicine) collected and compared information on known endocrine disrupting substances concerning structural alerts and (eco-)toxicological properties (*in vitro, in vivo*). The aim is to derive criteria for the identification of endocrine effect potential (preliminary indication) and priorisation based on the standard data set required by REACH.

necessary indicative endpoints in an fish screening assay to be able to detect also AR-interactions. For ER agonists and aromatase inhibitors, the available database and protocol is supposed to be sufficient to exclude false negative responses. Due to general problems with the interpretation of results concentrations > 0.1 * LC50 of acute fish tests should not be tested. The minimum of test concentrations should be three.

Tier 2 (issues 4-7). The population-relevant endpoints to be covered in a definitive test should be stage specific mortality (embryos, larvae, juveniles, adults), growth and development (time to hatch, hatching rate, length, weight, malformations, time to first spawn, sex ratio), general behavior and reproduction (fecundity, fertility, mating behavior). Sex ratio and time to first spawn are not included in the USEPA guideline. However, both endpoints are regarded important to determine the sensitivity of sexual endocrine effects.

Indicative endpoints such as Biomarkers (VTG, 11-kT), histological findings, secondary sex characteristics or behavior (mating, spawning) are not necessary for a definitive test as they are not appropriate to determine effects on the intrinsic growth rate of populations. However, the results of indicative endpoints can be used to exclude endocrine MoAs (e.g. liver toxicity identified by histopathology could be used to invalidate concern of sexual endocrine disruption raised by the 21d-FA) or to facilitate interpretation, e.g., concerning separation of primarily endocrine effects from general toxicity. Thus, the plausibility of regulatory NOECs for population-relevant endpoints can be supported.

Whether a FLCT or a 2-GT is the more appropriate definitive test could not be solved. Due to the strong commitment of the USA, the 2-GT will be internationally preferred. However, for reasons of animal welfare as well as of scientific plausibility (see 2.3.3.2), the 2-GT should be started with fertilized eggs of the F0 generation and ended with the determination of the sex ratio of the F1 generation. This was also discussed on an international expert meeting in Mallorca in September 2008 (ref. 26).

Abbreviated definitive tests at tier 2. For the reduction of regulatory efforts, a potential abbreviation of test protocols is of interest. However, there is need to demonstrate that the life stage most sensitive to exposure and the endpoints with most sensitive manifestation of endocrine effects are still covered. Thus, due to the limited database a decision to perform a FSDT or PLCT instead of a FLCT or 2-GT has to be well established.

The FSDT (OECD TG 2011) does not cover fecundity and fertility. Consequently, it is only suited as definitive test, if growth and sexual development are the most sensitive phases to a sexual endocrine MoA and relevant maternal transfer of effects is not probable. This was shown for aromatase inhibition as dominant MoA: for the relevant hazard assessments, the FSDT(-phase) was as sensitive as the FLCT. This seems also to be true for AR agonists. However, the database demonstrating sex ratio being the most sensitive endpoint is very limited. Potentially also for ER agonists the FSDT could replace the FLCT. The lower statistical power compared to fertility could be handled by using a higher uncertainty factor. Ongoing evaluation by industry and IME of existing FSDT and FLCT data on ER agonists aims at the establishment a sound comparison of sensitivities.

The PLCT or the short-term reproduction assay are not regarded appropriate as definitive tests, as the sexual development phase as most sensitive to exposure for the most sexual endocrine MoAs is not addressed.

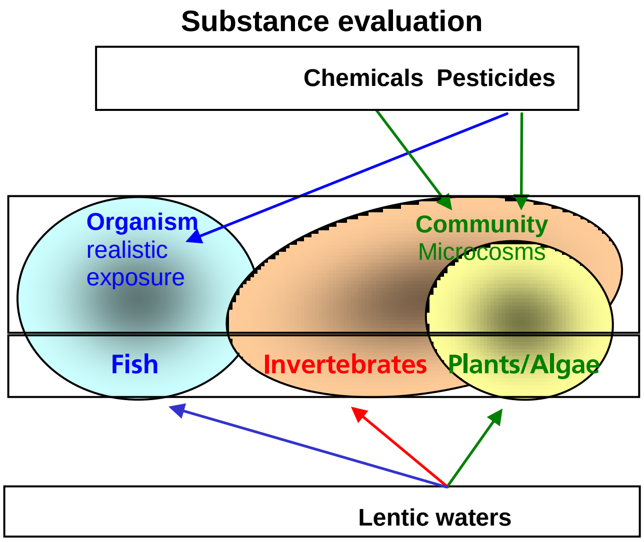
Conclusion. The actual data situation supports the proposed testing approach (Table 8) for ER agonists and aromatase inhibitors, when using the validated OECD test protocol for the first tier. For the other MoAs, the database is still too small to justify a comparable conclusion. The relevance of these other MoAs for the evaluation of sexual endocrine disruption potential of a wide range of substances is unclear, either. By theory and supported by few examples, ER antagonists and testosterone synthesis inhibitors should also be safely detected as sexual endocrine disrupters by the proposed testing scheme, if relevant (effect sufficiently below general toxicity). However, AR interactions cannot be safely detected by the proposed testing scheme, if validated OECD protocols are used, as VTG does not respond at relevant concentrations, secondary sex characteristics are not always sensitive enough and histopathology so far is not included. For AR antagonists, 11-kT seems to be a biomarker of appropriate sensitivity and independent of species differences. For AR agonists a sensitive biomarker still has to be identified.

The different safety of the proposed test strategy to detect different MoAs demonstrates the necessity to focus the test strategy right from the start. Thus, the basic tier at least has to specify whether there is concern about the androgen or the estrogen pathway to be able to include the appropriate endpoints in tier 1 testing. Alternatively, 11-kT and a specifically responding marker to AR agonists should be included in the OECD standard protocol, if available.

Outlook to other taxonomic classes. When evaluating endocrine disruption in aquatic systems, fish is of highest interest for humans, because as vertebrates, they are close to human physiology/toxicology and as food, they are most relevant for consumer protection. Other aquatic vertebrates are used for the investigation of thyroidal MoAs (amphibians). However, the most species-rich and ecologically important taxonomical classes of invertebrates are clearly under-represented in research, general knowledge on endocrine pathways and development of tests and test strategies (ref. 45). The IME agreed with Prof. Jörg Oehlmann, University of Frankfurt, to cooperate in GLP projects on snails and recently started an initiative for detecting ER interactions in annelids.

2.4 Ecological Interactions: Communities

2.4.1 Introduction



The most complex type of ecotoxicological studies is the investigation of realistic communities (Campbell et al. 1999). If well designed and conducted, these studies address all substance-specific concerns and uncertainty of the risk assessment is minimized (SANCO 2002). However, as variation of results is high due to conditions varying between replicates, biological variability and used methods of sampling and analysis, the identification and quantification of ecological effects is often imprecise. If no effect is visible at a realistic treatment, this may be sufficient for the evaluation of the specific use situation at the specific water body. It is not sufficient for an extrapolation to different use patterns, water qualities, and / or communities.

For an agreement on philosophies, methodologies, and techniques of so-called aquatic field tests, three international expert workshops were organized around 1990 in the US and Europe. Based on the recommendations, an OECD draft guidance document on freshwater lentic field test was prepared in 1996, mainly dealing with technical aspects of test performance. As European authorities were confronted with a growing number of mesocosm study reports in the following years, there was growing need for guidance regarding evaluation and interpretation of results within the scope of aquatic risk assessment, mainly for pesticides.

With growing experience in aquatic ecology and ecotoxicology it became more and more evident that even complex mesocosm studies have to be designed to address substance-specific concerns. This led to the development of indoor semi-realistic microcosms (as later defined by Campbell et al. 1999) during different projects financed by the German UBA, including guidance for focused community level studies (ref. 11) 52)). When initiating an UBA workshop on specific needs of aquatic risk assessment and resulting meso- and microcosm designs, it was realized that the time was obviously ripe for an international initiative. Theo Brock (Alterra, Wageningen, The Netherlands) had developed similar approaches and systems, the expert workshop HARAP (Higher tier Aquatic Risk Assessment for Pesticides, Campbell et all 1999) identified need for further guidance on “aquatic field studies”, and the OECD planned to modify and finalize the draft guidance document on freshwater lentic field tests (OECD 2006a).

In the following, the initiated international workshop (ref. 31) is shortly presented, as well as some microcosm studies addressing specific concerns of pesticide risk assessment or water quality objective setting. A comparison of the different approaches enables the derivation of general conclusions (or just experimental verification of known concepts) about community and population sensitivity depending on seasonal succession or nutrient status.

2.4.2 Community Level Aquatic System Studies Interpretation Criteria (CLASSIC)

From May 30 to June 2 1999, the workshop on Community Level Aquatic System Studies Interpretation Criteria (CLASSIC) was held at the Fraunhofer Institute, Schmallenberg, Germany. It was under the auspices of the European Union Commission (DG Health and Consumer Protection), the Organization for Economic Co-operation and Development (OECD), the Society of Environmental Toxicology and Chemistry (SETAC Europe), and the German Federal Biological Research Centre for Agriculture and Forestry (BBA) and Federal Environmental Agency (UBA). Objectives of the workshop were:

- To update, where appropriate, technical guidance on aquatic microcosm and mesocosm study protocols to include recent developments in the science and to provide guidance of the technical sufficiency of studies.
- To develop guidance on the ecological interpretation of aquatic microcosm and mesocosm data.
- To develop guidance on implementation of aquatic microcosm and mesocosm studies into regulatory risk assessment.

To address the objectives of the workshop, participants were assigned to a number of ‘break-out’ groups and provided with questions designed to stimulate discussions on topics associated with the subject areas outlined above. All workshop participants then discussed the findings of the break-out groups in plenary with the aim of arriving at a broad consensus. Reports of the breakout and plenary groups were appended to the consensus document, which thus contains the ‘raw data’ produced by the workshop. Subsequent to the workshop, these discussions have been developed into a concise guidance document by the workshop organizing committee (ref. 31). In summary, the 17 recommendations of this guidance are outlined below. Further discussion of each recommendation is included in the guidance document.

CLASSIC Recommendations:

1. An exposure-response experimental design with replication is clearly preferred. If possible, this should include the maximum predicted environmental concentration (PEC). The selected concentrations should generally be based on the expected effects and not only on the PEC. Where feasible, multiple applications should be avoided.

2. Previously, studies have generally attempted to simulate field exposure (“simulation” approach). Studies where the chemical is uniformly dosed into the water (“toxicological” approach) are also valid. They are often more easily interpreted and can be extrapolated to a variety of risk assessment scenarios.

3. Application of the test substance should be made in the period between spring and midsummer when the communities are in their ‘growth’ phases. Within this timeframe, species richness and abundance are usually most suitable, and the potential time available to observe recovery is long.

4. The level of taxonomic identification should be as high as scientifically justified or practically feasible (recognizing that there are constraints on species identification, especially for smaller species). Special efforts should be made for those groups that are identified in lower-tier studies as potentially the most sensitive.

5. Studying fish in micro- and mesocosms can present difficulties and needs to be carefully considered. When the invertebrate community is the principal endpoint of the study, it is recommended that free-living fish are not included.

6. Macrophytes are an important structural and functional component of shallow aquatic ecosystems, and in general should be included in micro- and mesocosm studies that aim to simulate these environments. If macrophyte communities are to be the principal endpoint of the study, special efforts are required to establish a diverse and representative community.

7. Univariate methods are recommended for investigating effects at the population level, and multivariate methods are recommended for describing community-level effects.
8. Ecological criteria should be taken into account when deriving acceptable concentrations from micro- and mesocosm data.
9. A framework for deriving ecologically acceptable concentrations (EACs) needs to be developed based on relevant endpoints from studies and considering other factors, which influence the dynamics of aquatic ecosystems.
10. Structural and functional endpoints are in general of the same importance. Species structure is usually the principal protection aim. Functional endpoints alone are not considered appropriate for protecting biodiversity.
11. If population recovery is demonstrated within an acceptable timeframe (which is species and context dependent), the observed initial effect should generally not be regarded as unacceptable, provided that adverse indirect effects do not occur.
12. If full recovery is not observed, it is recommended that additional tools (e.g. further laboratory studies) be used to address the remaining uncertainty. It is recommended that the use of ecological models for extrapolation be developed further in the future.
13. EACs from reliable static micro- or mesocosm studies should be regarded as generally representative or possibly conservative for surface waters in most agricultural landscapes.
14. Databases describing the abiotic and biotic conditions of surface water should be developed to aid interpretation and extrapolation between different waters and regions.
15. Landscape ecology should be considered when evaluating the uncertainty of mesocosm results because water bodies in agricultural landscapes are often not isolated and/or completely exposed.
16. When using mesocosm data in risk assessment, it is important to have transparent legal, political or societal protection aims (e.g., overall species diversity and ecosystem structure and function). Regulatory authorities should translate these aims into a form, which can be used to protect relevant aquatic ecosystems.
17. Additional guidance, training and tools are needed by those for conducting and evaluating microcosm or mesocosm studies.

Figure 18: Indoor microcosm facility for focused community level and extended laboratory studies Water and sediment from natural sources; Volume 16 * 1 m³; controlled temperature and light regimen (25 000 lx, sunlight spectrum); application of radio-labeled test substances possible

2.4.3 Development of an indoor microcosm system simulating seasons and climates

In the 1990s, an indoor microcosm facility was developed at the Fraunhofer IME, mainly for investigating fate aspects of radiolabeled pesticides. It consists of 16 vessels of approximately 1 m³ volume each, located on the floor of a greenhouse, partly containing aquatic lysimeter devices. The facility is temperature-controlled by evaporation coolers at 20 ± 5 °C, each vessel receiving light from a metal halide lamp (HQI TS 1000 W/S, Osram, München, Germany; wavelength area 280-780 nm, 25 000 lx at the water surface) Since 1996, they are mainly used for aquatic community level studies. In 2000, a second set of 16 glass microcosms was installed below the floor (Figure 18), receiving only light at the water surface (Osram HQI TS 1000 W/S). The systems can be observed from downstairs. The temperature is independently controlled by a heating and cooling liquid circulating in a stainless steel compartment forming the wall between two systems. Four vessels each can be controlled separately in a range between 5 and 40 ± 1°C. By controlling temperature and day length, different seasons and climates can be simulated in this worldwide unique facility. Both facilities are filled with water and sediment from natural sources that can be chosen individually for each study, depending on the needs concerning trophic level and species assemblage.

The microcosms system are used to derive a No Observed Ecologically Adverse Effect Concentration (NOEAEC, SANCO 2002), by identifying concentration-effect relationships for direct and indirect effects on the population and community level, and demonstrating recovery. An adverse effect would be an effect lasting until the end of the study. At the NOEAEC, occurring initial effects have to be completely recovered in accepted time intervals. For zooplanktonic communities, Brock et al. (2000) proposed a maximum of 8 weeks to be acceptable.

The investigated communities are no synthesized assemblages, but comprise of species introduced with unfiltered water and unsifted sediment. Vertebrates and macrophytes are excluded. The introduction of *Elodea densa* at one quarter of the sediment surface ensures system stability as well as space for planktonic organisms. The microcosms may be supplemented with additional species to ensure sufficient representation of species of concern. The statistical evaluation is based on univariate evaluation of sufficiently abundant species and on multivariate community analyses, i.e., diversity and similarity indices and principal response curves (PRCs, Van den Brink and Ter Braak 1999). Potential specific effects on benthic invertebrates or fish may be investigated by introducing encaged organisms of performing bioassays with water or sediment samples taken at appropriate sampling dates. The fate of the test substance may be investigated thoroughly by applying a sufficiently high concentration of radiolabeled material to a further microcosm. During the pre-treatment period the microcosms are connected to one another by a tube system (flow driven by aeration) to ensure similar conditions and colonization. The systems are separated directly before the first treatment. The exact water height at the start of the experiment is noted (the same heights ± 1 cm for each system) and used for calculations of the amount of test substance needed for application. Losses of water by evaporation are made up with de-ionized water if the water height decreases by ≥ 5%.

A study comprises of four to five concentrations in two or three replicates for effect testing, and one or two fate systems, commonly added by four controls. The test substance may be applied by spraying or mixed with slurry to simulate the most important route of entry to the water body (spray-drift or runoff; simulation approach), or by dosing it to the water column and stirring to achieve homogeneous distribution (toxicological approach). The test duration depend on the study objective. Successfully performed studies had a duration between 28 days (extended fish study (ref. 83)) and 13 months (long-term recovery study with persistent products (ref. 67)).

If communities of the region are included, the species diversity is low, as the mountain streams rather than lentic waters dominate landscape. The taxa usually identified in sufficient abundance for statistical evaluation are 11 zooplankton taxa (1 copepod, adults and nauplia; 1 ostracod (not differentiated to species), 4 cladocerans, 5 rotifers), 20-30 phytoplankton taxa (10-15 green algae, 2-3 cryptophytes, 3-4 blue-green algae, 2-4 species of other groups, diatoms only clustered in centrales and pennales, picoplankton < 5 µm not differentiated). The limited number of planktonic species is usually no shortcoming, because representation is sufficient and sensitivity is high du to the low variability and high statistical power. However, for macroinvertebates 5-10 taxa are clearly not sufficient. This has to be overcome by including sediment from other sources, introducing additional species, performing bioassays, addressing specific concerns by species sensitivity distributions (see chapter 2.2) or by performing outdoor studies.

2.4.4 Peak exposure: Pesticide registration studies

The primary contamination of surface waters with pesticides regarded in risk assessments is characterized by peak exposure of ditches, streams or ponds adjacent to agriculture, following spray drift, drainage or runoff during or after pesticide application. As pesticides are specifically designed with respect to fate and effect properties, also the side effects and concerns are specific, needing specifically designed studies to address them satisfactorily (Campbell et al. 1999).

At the same time, the EU guidance document on aquatic ecotoxicology (SANCO 2002) can be understood as clearly preferring outdoor mesocosm studies, regarding realism and species richness more important for the reduction of uncertainty than representation or statistical power. This is scientifically questionable. Anyway, as clients tend to prefer outdoor studies, in the last two years we built up close cooperation with gaiac, Aachen, and Mesocosm GmbH, Homberg/Ohm as principal investigators to perform outdoor studies in a multiple pond system and two enclosure systems, respectively (ref. 72)-75), 78)-80), 82)). The study design, the statistical and ecological evaluation and reporting was in the responsibility of Fraunhofer IME. In the following, only principles and general aspects of the indoor studies (ref. 67)-71), 76), 77), 81)) will be presented.

2.4.4.1 Herbicide studies

Herbicides are characterized by effects on photosynthesis, specific plant anabolism and plant growth. If acting via root uptake, they are designed to be water soluble and stable in aquatic solution, resulting in a high probability of entering water bodies also by runoff and drainage and of leaching to the groundwater. The communities being most at risk are phytoplankton, periphyton and macrophytes. Invertebrates may mainly be affected indirectly by a reduction of appropriate food. Fish sometimes accumulate herbicides or suffer from unforeseen side effects like endocrine disruption, which may be of interest due to high dissipation times from the water phase.

With respect to the concerns well addressed by community level studies, fish and macrophytes should be investigated in targeted laboratory studies (see chapters 2.3, 2.4.4.4 and 2.4.4.5). As the taxonomical differentiation of periphyton is difficult, this community is mostly covered only by functional endpoints or by thorough phytoplankton analyses regarded representative also for the periphyton species. Thus, community level studies with herbicides mostly focus on direct effects on phytoplankton and indirect effects on zooplankton including recovery. To demonstrate that this comparably simple task is not trivial, two problematic study results are presented, both leading to a repetition of the study in a more complex system.

Example 1: An indoor microcosm study was performed for generating data for the risk assessment of a critical active substance being used in many formulations and combinations with other active substances. In the most sensitive tier 1 algae study, a formulation with two active substances was tested, pointing to a synergistic effect. The lowbudget microcosm study (performed before CLASSIC), investigating only plankton organisms and *Lemna,* had several shortcomings:

- It was performed with a formulation only containing the most critical active ingredient to be generally used for risk assessment. However, the critical formulation was the economically most important one.
- The concentration range tested was oriented at the most sensitive effect concentration in the standard algae test and the PECs. It was intended to show that no effect would happen at the relevant PECs.
- The additional illumination seemed to be too intensive for *Lemna*. At the end of the study, *Lemna* seemed to be affected in all microcosms including the controls.
- As the sampling design was focused on phytoplankton, the volume was critical for investigation zooplankton, which appeared to exhibit low densities.

The study results indicated a NOEC at the highest test concentration. However, as no effect could be demonstrated, the authority could not find evidence that the study was designed properly to be able to detect sensitive effects. We tried to demonstrate a slight effect at the highest concentration in *Lemna*, measured in the first part of the study before decrease of the health status, but this was neither convincing, nor sufficient, as *Lemna,* if accepted as most sensitive species, was the only representative of aquatic macrophytes.

In the meantime, two complex outdoor mesocosm studies have been performed with different formulations containing the critical active substance (ref. 73), 74)). The highest concentration was three times the highest one of the indoor microcosm study. Different macrophytes and periphyton measurements were included as well as thorough investigations of indirect effects on zooplankton and macroinvertebates. The results were consistent with that of the indoor microcosm study: the threshold concentration of effects was close to the highest concentration in the indoor study, the most sensitive organism was *Lemna*. However, since a clear concentration-effect relationship has been demonstrated for all primary producing communities, the data will be regarded more reliable.

Example 2: An indoor microcosm study was performed with a focus on algae, as previous ecotoxicological testing identified green algae as the most sensitive aquatic species (ref. 68). Additionally, effects to functional endpoints such as periphyton and macrophyte production and water quality were investigated, as well as effects on zooplankton and benthic organisms. Six concentrations were tested in two replicates each, except the highest concentration which was only tested once.

In the course of the study, 20 different taxa of algae were determined. As a large number of taxa rarely occurred, in the controls an average of only four taxa was determined at each date of analysis. This was mainly due to the dominance of a single species, *Ankyra judayi*, which was a unique situation in all studies performed so far. A multivariate analysis of the phytoplankton community also included the rarely occurring species. Principal response curves showed clear effects until day 14; a recovery of the community until day 42 was observed for all concentrations (Figure 19). Stronger deviations from the control were found only for the microcosms treated a concentration of 0.45 µg/L.

The green alga *Ankyra judayi* was the dominating species in the pretreatment phase and showed the strongest effects (NOEC = 0.07 µg/L at day 3 and 5, Figure 20). Up to a concentration of 0.18 µg/L, total recovery was observed within 3 weeks. After approx. 5 to 6 weeks, almost total recovery could be observed even for the highest concentration (7 µg/L). At the end of the study, the highest abundances of *Ankyra* were found at concentrations of 7 µg/L. The significant recovery at 7 µg/L cannot be interpreted as an outlier, since the measurements clearly demonstrate that growth of *Ankyra* was possible at this concentration level. The microcosms treated with concentrations of 0.45 µg/L did not recover. However, this was not interpreted as caused by the test substance, since population growth was clearly observed at higher concentrations. One replicate of the controls also showed rather low abundances of *Ankyra* in the course of the study (after approx. 28 days). Therefore, the stop of *Ankyra* growth at 0.45 µg/L was considered as a randomly occurring effect or due to another cause of damage.

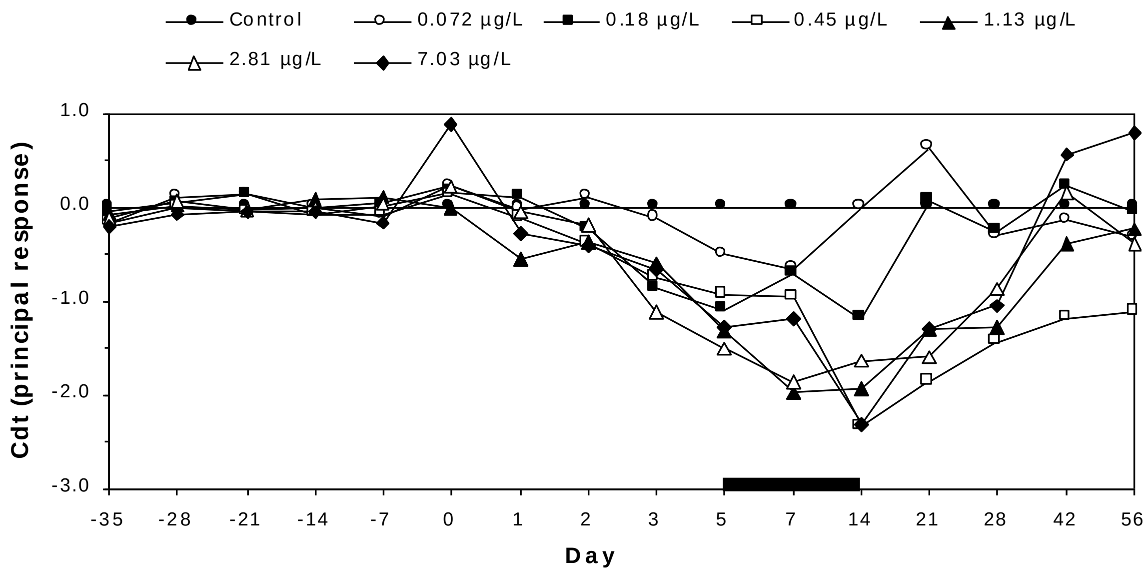


Figure 19: Principal Response Curves for the phytoplankton
The thick line indicates significant differences between the treatments.

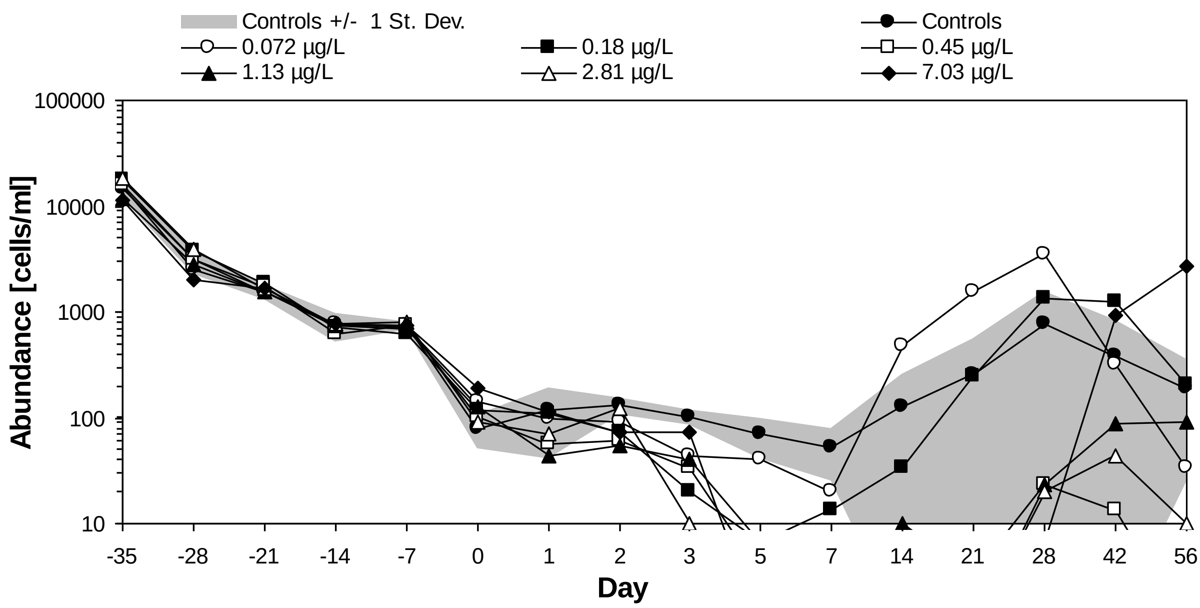


Figure 20: Population dynamics of *Ankyra judayi*: mean values for each treatment.

Results of a comparative laboratory study demonstrated that the species Ankyra, which dominated the microcosm study, was the most sensitive among the tested species, mainly due to its relatively slow growth rate. Therefore, it was suggested that there is no need to apply a safety factor covering the differences between species, or a very low factor could be used. The results demonstrate that up to the highest test concentrations of 2.8 and 7 µg/L the system is able to allow total recovery even for the most sensitive species of algae. Therefore, it was concluded that the concentrations were ecologically acceptable, especially when taking into consideration that significant impacts on functional parameters were not observed at the end of the study.

The authority did not follow the evaluation for three reasons:

- The number of sufficiently abundant algae species was regarded too low
- The effect at 0.45 µg/L did not recover in both replicates. This was not accepted as a random effect.
- The highest concentration tested as a unicate could not be included in the statistical evaluation.

Conclusions: A clear and consistent concentration-response relationship of direct effects on algae is a prerequisite for an acceptable risk-assessment of herbicides. Due to the worst-case exposure in laboratory standard tests, the sensitivity in community level studies may easily be overestimated. Thus, the choice of a high concentration is helpful. At the same time, this should be included in the statistical evaluation. Thus, at least two replicates of all concentrations are necessary.

Especially phytoplankton species and filamentous algae tend to develop extreme dominance structures that may corroborate biodiversity and recovery potentials. This phenomenon most easily occurs in small systems of low habitat complexity as represented by indoor microcosms. Consequently, especially for the smallest organisms as most dynamic ones, the bigger and more complex outdoor systems can be recommended. As a positive side effect, the relative sensitivity of macrophyte species can also be touched.

2.4.4.2 Insecticide and acaricide studies

Effects on arthropod physiology, mostly affecting the nerve systems, energy metabolism or reproduction, characterize insecticides and acaricides. Substances affecting signal transfer at synapses, as organophosphates and carbamates, they are water soluble, but often instable due to hydrolysis. The community being most at risk is zooplankton, especially high-performance filter feeders (chapter 2.2.3; Brock et al. 2000). If interacting with membranes as respiratory chain inhibitors, they are very adsorptive and may affect all invertebrates and fish. Pyrethroids mainly affect aquatic insects and higher crustaceans, such as amphipods, decapods and asselids (chapter 2.2.2).

Often fish are sensitive to insecticides, particularly rainbow trout. However, fish should not be included in community level studies used for invertebrate evaluation. They have to be addressed separately. If there is need, Fraunhofer IME has a developed a special study design for rainbow trout extended 28 d studies in microcosms (chapter 2.4.4.4). For aquatic invertebrates, next to further laboratory tests with other test species and laboratory studies under more realistic exposure conditions, a micro- or mesocosm study can be performed. However, the exposure of aquatic invertebrates to the test substance should be representative, i.e. simulating the relevant entry routes into water bodies. The study should evaluate representatives of all sensitive immanent freshwater invertebrate species.

If the known background information on sensitive species justifies focusing the study on planktonic species and benthic crustaceans, emergence has not necessarily to be investigated; aquatic insects need not to be focused on. This justifies the performance either as outdoor or as indoor study. The advantages of indoor and outdoor studies can be summarized as shown in Table 9. If aquatic insects and emergence have to be focused on, indoor studies are the worse option, as recovery is difficult and the sediment surface area may be too small.

Table 9: Advantages of indoor and outdoor community level studies

	Advantages of indoor studies	Advantages of outdoor studies
Species richness	-	+
Data variability	+	O/-
Management	+	O
Environmental realism	-	+
Worst case exposure	+	O
Fate studies	+	O
Costs	+	-

As follows, the acceptability of the study results by authorities differ depending on the philosophy: Authorities with focus on environmental realism and species richness (BVL) clearly favor outdoor studies, authorities focusing on precision of results and the precautionary principle (UBA) favor indoor studies if properly designed. To present some specific aspects of insecticide studies performed in indoor microcosms, three examples of zooplankton community studies are elucidated.

When predicting pesticide concentrations and effects in surface waters, advanced available models include the volume of a water body in a specific scenario and dilution by the flow rate in lotic waters. The highest loadings result from spray drift to shallow ponds or lentic ditches after pesticide application in orchards. In this specific worst case situation, the applied pesticides settle on the water surface, forming hydrophobic layers, and distribute to the deeper parts of the water body depending on formulation properties such as dispersion or emulsion capability, and active substance properties, such as photolysis potential or adsorption to macrophytes. The resulting potential stratification of concentrations in the water body is not addressed by exposure models. Complex regulatory effect studies such as mesocosm studies are designed to gather integrative and representative total effect data for each taxon of interest. They consider vertical migration of zooplankton populations by depth-integrated sampling, and thus intend to level out dynamic processes rather than to identify or quantify them. Circadian rhythms of planktonic stratification due to light regimens and predator avoidance are well described. However, thorough investigations on stratification of abundances due to chemical stress avoidance are missing.

Example 1: As a higher-tier study for the aquatic risk assessment of the test substance, an acaricide with a dominant acute mode of action (respiratory electron transport inhibition), a 4-week microcosm study was performed focusing on zooplankton and excluding fish (ref. 19). The test substance is photodegradable and adsorptive.

Materials and methods

Seven microcosm tanks were filled with a layer of silty sediment of 20 (± 2) cm height and an overlying water body of 60 (± 1) cm height (= 600 L). Water and sediment were taken from natural sources containing the indigenous communities, not filtered or sieved. Additionally introduced macrophytes (*Elodea sp.*) served as habitat structure and provided functional stability by photosynthesis as well as enhanced surface area or adsorption. The macrophytes did not occupy more than one quarter of the water column. The microcosms were installed in a temperature-controlled greenhouse at 14 ± 3 °C. Metal halide lamps with mean light intensities at the water surface center being approximately 25 klx provided additional illumination (16/8 h light/dark rhythm for simulation of summer irradiation).

During the 12-week pre-treatment period, the microcosms were connected to one another by a tube system ensuring a slight water exchange between the systems, driven by aeration. At test start, the microcosms were disconnected and treated with test substance by spraying the formulated test substance (“Kiron”) on the water surface, simulating spray drift under field conditions. Sufficient stability of Test substance in the application solutions was proven by chemical analyses of the application solutions after preparation and directly before application. Six of the seven microcosms were treated with nominal 0.10, 0.32, 0.56, 1.0, 3.2 and 10 µg active substance / L. One microcosm was not treated and served as a control.

The study focused on effects at a mean water depth of 30 cm during the first days of exposure to be comparable to an extended fish study with the same exposure and fish encaged at a water depth between 15 and 45 cm. Test substance concentrations were measured at a depth of 30 cm and additionally near the water surface at 10 cm depth. Water samples were taken by means of thin glass tubes that were introduced in the water phase cautiously in order to minimize disturbance. On each sampling date, six times 100 mL of water (randomized locations) were sampled from each microcosm at both water depths. The six samples from each depth were pooled and repeatedly extracted with CH_2Cl_2. The CH_2Cl_2 phases were combined, reduced to dryness, re-dissolved in 1 mL of $CH_3OH/H_2O/HCOOH$ 90:10:8 (v/v/v) and centrifuged at 11000 g for the separation of suspended particles. The Test substance concentration was measured using high-performance liquid chromatography / tandem mass spectrometry (HPLC/MS/MS) at a LOQ of 0.01 µg Test substance/L water sample.

Zooplankton were sampled weekly in the pre-treatment phase, on day 0 (before application), and on days 1, 2, 4, 7, 14, 21, and 28 after application by pumping 5 L (before treatment) or 2 x 5 L (after treatment) of water through a steel tube which was carefully introduced into the water column. During the sampling period, the steel tube was only moved slowly in a vertical direction between about 15 and 45 cm of the water column. From day 0 on, the sampled water was not returned to the microcosms in order not to disturb the water column. Thus, each sampling resulted in a decrease of the water depth of about 1 cm. Each sample was sieved (15 µm mesh width), re-suspended in 50 mL, and conserved with Lugol’s solution (1%). After sedimentation, the sample volume was reduced to 2 mL. The organisms were equally distributed and an aliquot of 0.5 mL was examined. The densities were expressed as abundances per L water volume. Taxonomic analyses were conducted in two tiers: first to the Family level, then rotifers, copepods and cladocerans were differentiated to species or to the lowest taxonomic level possible.

The test design allowed the calculation of acute effect concentrations by using regression analysis. Zooplankton abundances were related to mean abundances of the respective system during the last days of the pre-treatment period and calculated relative to the respective value of the control system. As the interconnections during the pre-treatment period were only able to provide all systems with the same plankton taxa for colonization, a certain between-microcosm variability regarding abundances and community composition remained. Therefore, a relation to the mean abundances of the pre-treatment period seems to be justified.

Results

The application solutions were analyzed for the Test substance concentrations, confirming the nominal loading. The sprayed test item formed a film on the water surface with slow distribution into the water column. After 24 h, the highest concentrations were measured with approximately 50% of nominal at a depth of 10 cm but assumed to have occurred earlier. At that time, at a water depth of 30 cm the concentrations were already slightly higher (approx. 60%) than at 10 cm depth, which might be due to photodegradation in the upper part of the water column, as photodegradability is a relevant test substance property. Thereafter, the concentrations decreased due to distribution to deeper parts of the water column and the sediment as well as due to degradation. At a water depth of 30 cm, Test substance disappeared more slowly than at a water depth of 10 cm, especially at the higher concentrations. At the end of the study, 28 d after treatment, 5% of nominal concentrations were measured at 10 cm depth and 12 – 15% of nominal at 30 cm depth.

Looking at the total zooplanktonic community, the major parts being copepods and cladocerans, a more or less stable oscillation of all the systems were observed during the pre-treatment phase as well as during the post-treatment period for the control and the lower concentrations. The range was approximately 10 to 100 individuals per liter (maximal difference between systems: factor of 5). Regarding treatment effects, two different types of effects can be distinguished: short-term effects followed by rapid recovery within one day and long-term effects with a trend to but not full recovery after 4 weeks at the highest test concentration.

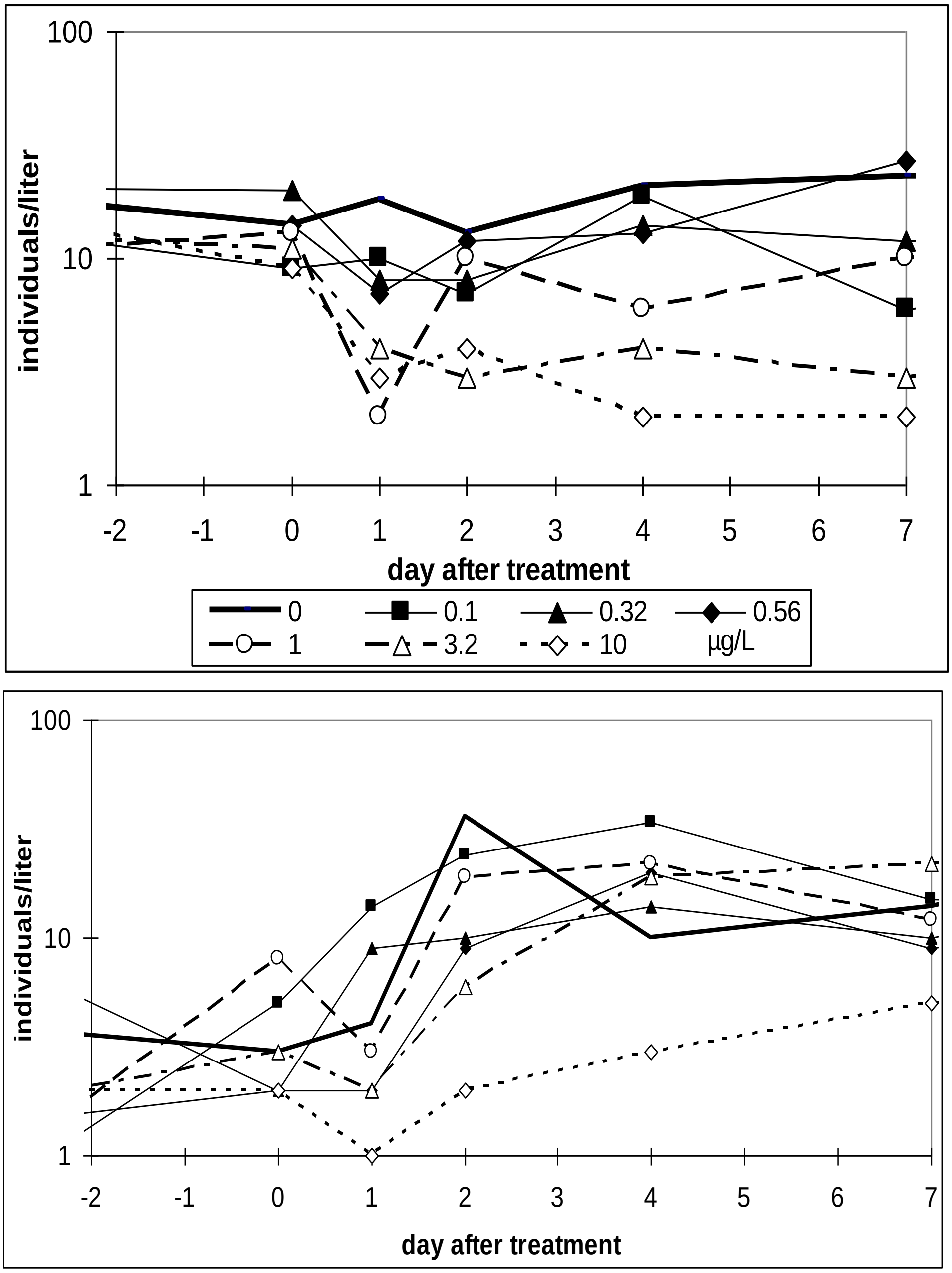


Figure 21: Acute effects of test substance on zooplankton;
upper: Cyclopoid copepods, lower: Cladocerans

Copepods (*Mesocyclops sp*) exhibited a clear concentration-response relationship. When including the starting point of the populations on day 0, the EC50 was calculated to be 0.28 µg/L. A clear reduction of the population density at the sampled depth started at 0.32 µg/L (Figure 21). The population density was fully restored up to 1 µg/L one day later. Population effects were observed at 3.2 µg/L (full recovery within four weeks) and 10 µg/L (no full recovery after 4 weeks). The effects were similar in adults (10-20% of the sampled copepods) and juvenile stages (80-90%).

Cladocerans (*Daphnia magna, D. longispina, D. pulex, Acroperus harpae, Chydorus sphaericus*) reacted less clearly. The EC50 was calculated as approximately 3 µg/L, with reductions in populations densities starting at 1.0 µg/L. The population density was fully restored up to 3.2 µg/L 1-3 days later (Figure 21). Population effects were observed at 10 µg/L (full recovery within four weeks).

For ostracods (not differentiated) and rotifers (three taxa), the low and discontinuous abundances in the samples did not enable calculation of any effect concentration. All taxa observed in the control during the exposure period were also found up to the highest test concentration. When looking at phytoplankton primary producers, the functional parameters of photosynthetic pigment concentrations in the water gave no indication for direct or indirect effects of the treatment on algae. No clear concentration-response relationship was observed concerning a shift in the relations between the different pigments.

Discussion and conclusions

The results indicate that the temporary reduction of population density was due to flight reactions at concentrations up to 1.0 µg/L (copepods including nauplia) or 3.2 µg/L (cladocerans) at the sampled water depth. The sinking front of the test substance in the water column may have repelled the zooplankton, which seemed to escape into the deeper layers of the water column. When the test item reached a more homogeneous distribution, the populations re-settled, the time being too short for a recovery by reproduction. Thus, zooplanktonic crustaceans have to be able to detect exposure gradients and react appropriately without apparent adverse effects on individuals. This ability was more pronounced in the copepods of all life stages than in cladocerans. It is consistent with descriptions of zooplankton flight reactions from predators (Drenner and McComas 1980; Winfield et al. 1983), showing that flight is a usual reaction of planktonic copepods, whereas it is less effective in cladocerans or rotifers. The difference between the taxonomic groups might be due to better chemical sensoric capabilities of copepods (Butler et al. 1989). Assuming that crustacean early life stages with frequent molting are more sensitive than adults (Hanazato and Hirokawa 2004), detectable effects being more pronounced in copepods compared to cladocerans could also be explained by population structure. Whereas copepod populations in the water column mainly consist of nauplia, followed by copepodits and adults (mean numbers of nauplia in the present study: 80-90%), cladoceran populations near the state of density equilibrium tend to have a significantly smaller proportion of early life stages (below 10%, Hammers-Wirtz and Ratte 2004).

Flight reactions from chemical stress were addressed by a novel approach in earthworm testing alternatives (Hund-Rinke et al. 2005), showing that annelids also are able to detect and effectively avoid different harmful stressors. The nature of sensory cues, including direct sublethal stress, resulting in behavioral action in invertebrates, remains unclear and needs targeted investigations.

If active avoidance of chemical stress can be regarded as relevant for zooplanktonic crustaceans including their juvenile stages, it may be a considerable mitigation measure in structured habitats along moving exposure gradients and has to be taken into account when interpreting post application monitoring results (see chapter 3). Müller (pers. commun.) repeatedly found decreases of copepod abundances in the field to be the most sensitive effect of insecticides, which is not supported by laboratory tests but may be explained by the avoidance hypothesis. We investigated periphyton-related benthic organisms in a recirculating outdoor artificial stream system after a local overspray application of an adsorptive acaricide in autumn[5] (see chapter 3.5). Abundances of cyclopoid copepods and chironomids as the most mobile organisms were found to be the most sensitive endpoints directly downstream of the application, but showed a fast recovery in spite of the low temperatures, indicating that re-colonization by escaped organisms might have been more important than recovery by reproduction.

Any discussion on potential implications for the conduct of regulatory micro- or mesocosm studies should focus on test substance properties and application scenarios. Giddings et al. (2002) mentioned the possibilities of a "toxicological approach" versus a "simulation approach". A toxicological approach tries to generate a homogeneous distribution of the test substance in the water column. The test results are less dependent on the test scenario and can be extrapolated more easily. By using this approach, avoidance reactions are not possible. However, testing substances that are highly adsorptive, a toxicological approach is neither realistic nor feasible, as the equal distribution would result in unrealistic losses of the test substance from the water column to macrophytes and sediment by enhancing the contact probability in time. Photodegradation may be more important for hydrophobic films at the water surface compared to homogeneously dissolved test substance concentrations. Thus, for application scenarios resulting in entrance routes dominated by spray drift and using adsorptive and photodegradable test substances, a simulation approach, applying the tested formulation by overspray to the water surface, seems to be more appropriate. In such studies, a location- and depth-specific sampling may provide the best data for a scientific investigation of avoidance and re-colonization, being necessary for a detailed extrapolation to field situations.

Example 2. The effects of an acetylcholinesterase-inhibitor on the community of indoor aquatic microcosms (ref. 69) were investigated after of acute invertebrate toxicity had shown that cladocerans were the most sensitive species. For sensitive testing of realistic acute effects on cladoceran populations as well as recovery by reproduction, indoor systems are well suited. The occurrence of and time needed for recovery was recorded to derive an EAC, now reduced to a NOEAEC (SANCO 2002). An effect lasting until the end of the study was defined an adverse effect. The investigated zooplankton community comprised of species introduced with natural water and sediment (organic carbon content of 3%). The statistical evaluation was based on sufficiently abundant zooplankton species and on multi-variate community analyses. Beside the investigation of species-specific abundances of zooplankton, sampled by integrating all depths of the water body, the study included an assessment of indirect effects on further parts of the aquatic community (phytoplankton, macrophytes) which were measured as functional parameters, such as particle density and photosynthetic pigment concentration in water or as macrophyte biomass production. The test substance was tested at nominal initial concentrations of 0.25, 0.8, 2.5, 8.0, 25 and 80 µg/L in two replicates each (four untreated controls), applied via spraying on the water surface two times with an interval of 50 days. The endpoints were observed for 115 days in order to determine the times needed for recovery. The recovery potential due to decreasing test substance concentrations in the water column was tested by accompanying *Daphnia magna* bioassays with microcosm water sampled at different times after the treatments.

The initial concentrations were calculated from measured concentrations of application solutions and confirmed by calculating t=0 values when fitting simple first order e-functions to measured phosalone concentrations of microcosm surface water. Thus, effects could be

[5] Schäfers C, Schmitz A, and Hassink J. 1998. Possibilities and limitations of simulating fate and effects of pollutants in a stream mesocosm. Platform presentation, 8th Annual Meeting of SETAC Europe, Bordeaux, France, April 14-17, 1998

related to nominal initial concentrations. The concentration of the test substance decreased with a half life in water of 1.9 days at 16 to 20°C during 10 days after the first application, and with a half life in water of 1.2 days at 20 to 21°C after the second application.

The test substance caused clear direct effects reducing zooplankton populations, especially cladocerans, resulting in algal blooms as indirect effects. The zooplankton community NOEC obtained by multivariate analysis was determined to be 2.5 µg/L, whereas for the three main cladoceran populations it was 0.8 µg/L on days within the first week after the treatments. The lowest daphnid EC50 values were calculated with 1 to 2 µg/L.

Additional laboratory bioassays with microcosm water sampled at different times and *Daphnia magna* produced comparable results when related to actual concentrations of the microcosms. The mean EC50 value of 0.33 µg/L corresponded well to the result of a Daphnia acute immobilization test according to OECD TG 202 (0.40 µg/L). When related to nominal = initial concentrations, the exponential doubling time of the EC50 was calculated to be 1.5 and 1.3 after the first and second, treatment, respectively. These correspond to the test substance half-life, the exponential fit being even better than that to the analytical data.

In the zooplankton community, rotifers and adult copepods profited from the reduction of cladocerans and exhibited population growth up to the highest concentrations. This was also evident for copepod nauplia at medium concentrations (2.5 and 8 µg/L). At 25 and 80 µg/L, however, they showed decreasing populations within the first two weeks after the first treatment. Ostracods remained indifferent and did not show effects.

Indirect effects on phytoplankton were measured by means of particle counts, pigment analysis, and flow-cytometry. The abundance of particles below 2.5 µm diameter re-presented the most sensitive parameter (NOEC = 0.8 µg/L) within the first week after the second treatment. The index “phyllopods per million particles” even revealed trends of temporarily reduced feeding activity down to the lowest initial concentration tested (0.25 µg/L), being as sensitive as laboratory toxicity tests. Primary production, as the basic functional parameter of lentic systems, measured by the macrophyte production as well as by oxygen contents and pH, did not show any effect.

In conclusion, a very consistent picture of direct impacts causing indirect effects was derived over the full range of concentrations tested.

When looking at recovery times, a clear relationship to treatment levels was observed. All significant effects and trends that were observed up to initially 2.5 µg/L fully recovered within two weeks after the last treatment indicating a re-population from living stocks. All effects observed at the high initial concentrations of 25 and 80 µg/L fully recovered within six to nine weeks after the last treatment regardless of the definite concentration, indicating autochthonous re-population of all three cladoceran species from resting stages (ephippiae).

For the first two weeks after treatment, the single species NOEC for the most sensitive species was calculated to be 0.8 µg/L, and the zooplankton community NOEC to be 2.5 µg/L. At study termination nine weeks after the last treatment, neither a direct nor an indirect effect could be observed up to the highest concentration of 80 µg/L any longer, indicating full recovery.

When comparing the different univariate and multivariate endpoints due to derived NOEC and NOEAEC values (Table 10), the multivariate analysis was slightly more sensitive than the different community indices (e.g., Figure 22) regarding the NOEC. It was more consistent in concentration-response relationship after the first treatment and delivered a clearly different image of effects and recovery after the second treatment (Figure 23). Obviously, redundancy analysis is more robust towards increasing control variability compared to the community indices. Besides this, the PRCs provide valuable additional information on effect characteristics of the included species. Although in the end full recovery is observed by all evaluation methods, the result of the multivariate analysis is more reliable, as effects are not masked by control variability. The multivariate community analysis is as sensitive as the univariate analysis of the most sensitive taxon. This is because all three cladoceran species were comparably sensitive and dominated the species weights.

A main criticism at the study was the high pH of approximately 10 during exposure, causing hydrolysis of the test substance. The high pH was caused by the system photosynthesis (oxygen saturation of approximately 150%). This is a normal situation of small, macrophyte dominated water bodies, but certainly not a worst-case exposure condition.

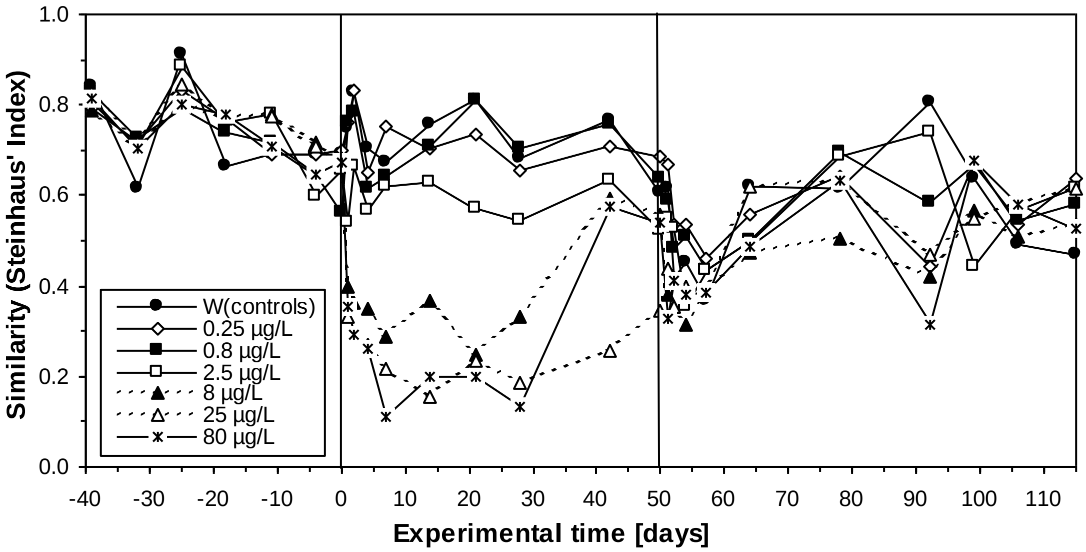


Figure 22: Similarity analysis using Steinhaus' index for the zooplankton
Similarities between controls and treatments respectively within controls.
Vertical lines indicate days of the two applications. Due to enhanced control variability, the effects after the second treatment are less pronounced.

Table 10: Comparison of NOEC and NOEAEC (according to Brock et al. 2000) values calculated by using different zooplankton community evaluation methods

Evaluation method	**NOEC**	**Full recovery within 8 weeks (NOEAEC)**	
		After 1st treatment (d 50)	**After 2nd treatment (d 106)**
Number of taxa	2.5 µg/L	80 µg/L	80 µg/L
Diversity (Shannon-index)	2.5 µg/L	80 µg/L	80 µg/L
Similarity (Steinhaus-index)	2.5 µg/L	80 µg/L	80 µg/L
Similarity (Stander's index)	2.5 µg/L	80 µg/L	80 µg/L
Redundancy analysis (PRC) PCA 1st axis sample scores	0.8 µg/L	8 µg/L	80 µg/L
Most sensitive taxon	0.8 µg/L	8 µg/L	80 µg/L

Note: In all cases hypothesis testing was performed by using Williams' test

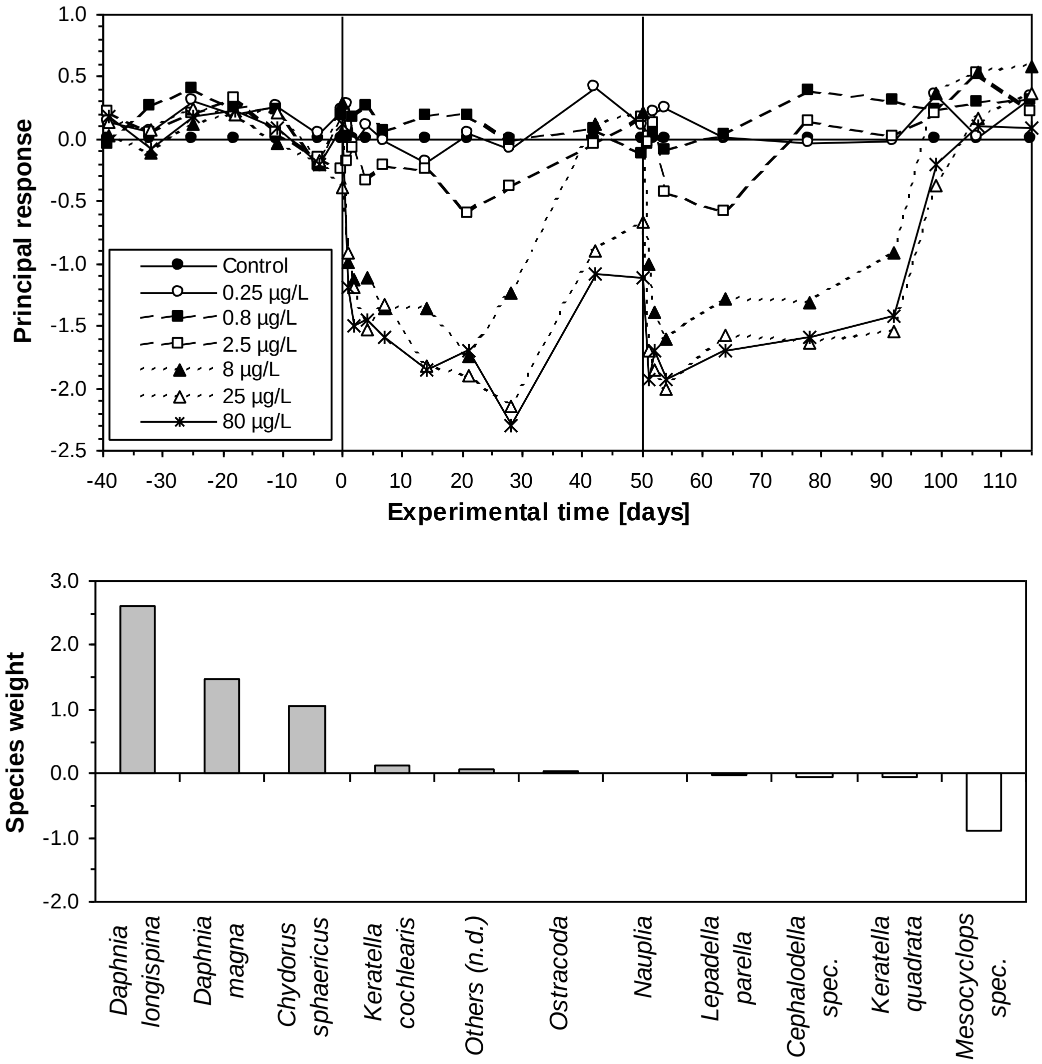


Figure 23: PRCs for the zooplankton (top) and related species weights (bottom)
PRCs indicate a significant part of total variance (p=0.005). Vertical lines indicate days of application.

Example 3. The lower tier data situation of an acetylcholinesterase-inhibitor leads to a request of a community level study, focusing on zooplankton comparably to example 2. The test substance, however, had a much higher hydrolytic instability. To be at least able to measure the test substance during the first day of exposure, a zooplankton community study was planned in indoor systems, which could be highly managed (ref. 70). To keep down the pH, the microcosms were aerated during the pre-treatment period, and the macrophytes (*Elodea densa*) were cut several times, the last time 11 days before treatment to reduce the phytosynthetically active biomass. By this measure, the pH was kept at approximately 8 at disconnection of the systems and treatment. Because of the reduction of the individual system size from 16 m³ to 16 * 1 m³ and the stop of water movement and aeration, the daily amplitudes in oxygen and CO_2-concentrations increased, resulting in increasing and very variable daily pH maxima. However, we were able to measure the test substance concentrations at six to eight sampling times (up to day four at the highest treatment) and could calculate a DT50.

The product containing the test substance was applied once by spraying on the water surface at nominal initial concentrations of 7, 23, 70, 225, 703 and 2250 µg test substance/L in two replicate systems each with three untreated control microcosms. The impact of the test substance on the zooplankton community introduced with water and sediment (organic carbon content: 4.5%) was observed for 56 days in order to determine the time to achieve recovery.

Beside the investigation of species-specific abundances of zooplankton, the study included an assessment of indirect effects on further parts of the aquatic community (phytoplankton, macrophytes), which were measured as functional parameters, such as particle density and photosynthetic pigment concentration in water or macrophyte biomass production.

The initial test substance concentrations calculated from the measured application solutions were in the range between ± 20% of the nominal values. After application the concentration of the test substance decreased with a half-life time of 5.1 hours (pH 7.9 to 8.5; temperature: 18 to 20°C).

The application of the test substance caused effects on the zooplankton taxa starting at a treatment level of 70 µg/L by reducing zooplankton populations, especially cladocerans, resulting in higher algal densities as indirect effects. The zooplankton community NOEC (based on abundances over the whole test duration) obtained by multivariate analysis was determined to be 23 µg/L. For the most sensitive cladoceran species found in the test systems, *Daphnia longispina and Daphnia magna*, the NOEC was also 23 µg/L, starting on day 4 and lasting until day 21 after the application.

Rotifers, adult copepods and nauplii showed short-term effects at the highest treatment level (2250 µg/L) only. At lower doses, these species had a benefit from the reduction of cladocerans and exhibited population growth. Ostracods were not affected in any of the test concentration levels.

Indirect effects on phytoplankton were measured by means of particle counts and pigment analysis. Particle counts revealed significant increases at a treatment level of 703 µg/L within the first week after the treatment. In the second week a NOEC for increasing densities was found to be 23 µg/L (particle 2.5 – 15 µm) corresponding to the NOEC for the decrease of Daphnia abundance. The primary production, as the basic functional parameter of lentic systems, measured by the macrophyte production as well as by oxygen contents and pH, did not show any effect up to the highest concentration levels.

In conclusion, at about three weeks after application a consistent picture of direct impacts causing indirect effects was derived for test concentrations of 70 µg/L and higher. Regarding the very fast dissipation of the test substance from the water column, the effects on zooplankton species manifested considerably late, especially when comparing the generally fast zooplankton response to direct effects as demonstrated in examples 1 and 2. The late effect was most probably caused by the primary metabolite of the test substance that still has acetylcholinestrase-inhibiting properties and is more persistent than the parent compound.

When comparing the different univariate and multivariate endpoints due to derived NOEC and NOEAEC values (Table 11), all analyses resulted in the same sensitivity regarding the NOEC, except the number of taxa and diversity. As the Shannon index tends to increase with an impact on dominant species, it cannot be considered as a sensitive parameter *per se.* With respect to recovery times, a clear relationship to treatment levels was observed. All endpoints with significant effects and trends, which could be observed up to the treatment level of 703 µg/L, fully recovered within the post treatment period of 8 weeks, indicating a re-population from living stocks. At the highest treatment of 2250 µg/L, *Daphnia longispina* showed a clear trend of recovery towards the end of the post-treatment period, but did not reach the level of the controls 56 days after application.

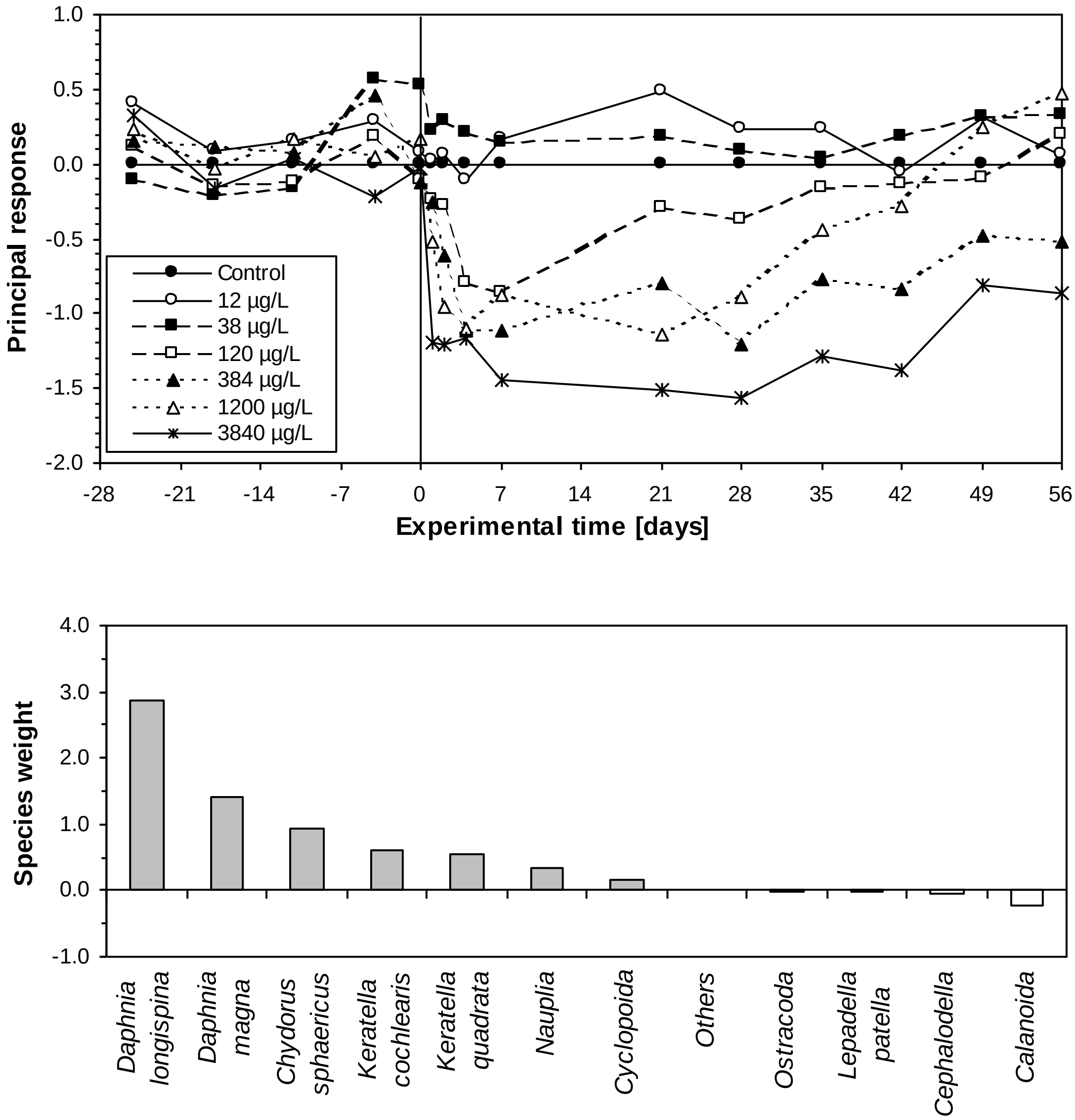


Figure 24: PRCs for the zooplankton (top) and related species weights (bottom)
PRCs indicate a significant part of total variance (p=0.005). The vertical line indicates the day of application.

In this case, the multivariate community analysis appears to be less sensitive than the univariate analysis of the most sensitive taxon. However, the Principal Response Curves (Figure 24) indicate a lasting community effect for the highest treatment level on day 56 after application, which lacks statistical significance due to replicate variability. The deviation of concentration 4 can be explained by the outlying character of one of its replicates. The particle counts and chlorophyll measurements consistently demonstrate an extraordinarily poor primary production in that replicate during the entire study, also resulting in low zooplankton community data. Therefore, the No Observed Ecologically Adverse Effect Concentration (NOEAEC) can be derived to be 703 µg test substance/L based on the demonstrated recovery of the most sensitive taxon in the study and the visual evaluation of the PRCs.

Table 11: Comparison of NOEC and NOEAEC (according to Brock et al. 2000) values calculated by using different zooplankton community evaluation methods

Evaluation method	NOEC	Full recovery within 8 weeks (NOEAEC)
Number of taxa	70 µg/L	2250 µg/L
Diversity (Shannon-index)	2250 µg/L	2250 µg/L
Similarity (Steinhaus-index)	23 µg/L	2250 µg/L
Similarity (Stander's index)	23 µg/L	2250 µg/L
Redundancy analysis (PRC) PCA 1st axis sample scores	23 µg/L	2250 µg/L
Most sensitive taxon	23 µg/L	703 µg/L

Note: In all cases hypothesis testing was performed by using Williams' test

2.4.4.3 Fungicide studies

Fungicides are characterized mainly by effects on growth metabolism, aiming at the inhibition of fast growing cells. Thus, they represent various modes of action, being mostly designed to be sufficiently persistent for preventing the expansion of an infestation by fungi. Consequently, fungicides are the most variable group of pesticides and communities being generally at a specific risk cannot be identified. A mesocosm study should represent all aquatic communities. Azole fungicides, inhibiting the fungi ergosterol synthesis, as a side effect are also able to inhibit the synthesis of steroid sex hormones by affecting the aromatase (17-beta estradiol synthesis), the 11-hydroxylase (11 keto-testosterone synthesis), or both. This could cause endocrine disruption in fish, which should be addressed in life cycle studies (see chapter 2.3). Similarly, specific concerns for macrophytes should be addressed in targeted laboratory studies (chapter 2.4.4.5).

As an extreme example of a fungicide study, a study with a copper hydroxide product is presented (ref. 67). In this study, the major aquatic community assemblages were represented (phyto- and zooplankton, macrophytes, benthic invertebrates). The application pattern reflected a fungicide multiple application scheme (spraying on the water surface six times with intervals of 10 days) and the duration of the study was more than one year (385 days) due to the persistence of the active substance, requiring an investigation of recovery until the following vegetation period. The specific considerations of metal toxicity and bioavailability are also reflected in the chapters 2.4.5 and 2.4.6.

Test conditions. The test substance was repeatedly applied by spraying on the water surface at nominal initial concentrations of 2.5, 12, 24, 120 and 240 µg copper/L in two replicate microcosms, each, with four untreated controls. The nominal concentrations reflect a range of concentrations starting from that resulting from worst-case use up to the nearly 100fold concentration to cover long-term effects of copper and recovery, as the persistence has to be specifically regarded in the risk assessment. Two further microcosms were used to investigate fate of copper in more detail, one applied with a high dose and one serving as background control. In the fate microcosms, sediment samples were investigated for AVS and redox potential 10 weeks after the last treatment and at the end of the study. The upper 5 cm of the sediment were sampled five times during the study, separated in sediment and interstitial water and both analyzed for total and dissolved (0.45 µm filtered) copper. At study termination, a copper balance was tried to determine by measuring water (3 depths), sediment (upper 5 cm layer and lower layer), and macrophytes, at the same time measuring or assessing dry weights as exactly as possible.

The investigated aquatic community comprised of species introduced with water and sediment (organic carbon content: 4.5%). No vertebrates were present. Indigenous Macrophytes were excluded as far as possible to minimize variability. Instead, 40 individual plants of *Elodea densa* were introduced on one quarter of the sediment area and managed by cutting to maintain three quarters of the water column for planktonic organisms. Investigated endpoints were the abundances of zoo- and phytoplankton species in the water column (depth integrated sampling), the growth of introduced macrophytes, the invertebrates found in sediment samples and emergence traps and the emergence of temporarily introduced chironomid larvae in specific in situ-tests.

The bioaccumulation of copper in benthic organisms was investigated by introducing snails to the test systems. Twenty adult snails from laboratory cultures were introduced to each system within the pre-treatment phase. The species was *Planorbarius corneus* (origin: ECT laboratories, Flörsheim, Germany), a planorbid snail, occurring naturally in small ponds in middle Europe. This group of snails has haemoglobin instead of haemocyanin as respiratory pigment and therefore should have a less specialized copper regulating metabolism. Prior to introduction, an acute range finding test with the test substance showed 100% mortality after three days at 2.4 mg/L (10-fold the highest concentration) and 25% mortality after 5 days at 240 µg/L (highest target concentration). Therefore, the fate system was treated with the second concentration of the test substance (120 µg/L). The accumulation of copper was also investigated in Macrophytes (using the biomass harvested for growth control) and planktonic organism.

The water temperature and light regimen should be similar to outdoor conditions throughout the year. Thus, the periods of additional illumination were adapted to the yearly light regimen of the geographical latitude in terms of day length (natural photoperiod was supported by providing light intensities and spectral qualities of a cloudy day). Temperature regulation was limited to the aeration and evaporation cooling systems of the greenhouse facility. The measured differences of individual microcosms of up to 2 °C were negligible compared to daily differences, differences between days and amplitudes throughout the year. The dynamics over the year was roughly comparable to the mean temperature regime of German sites (Figure 25). Compared with mean air temperatures of Karlsruhe (1960-1990),

- the amplitudes of water temperature should be more moderate. For example, lowest values should be around 4 °C (lowest temperature of deeper water bodies). This criterion was met by the study conditions. Normally the uppermost water temperature values should be lower than the corresponding air temperatures. The fact that this was reversed was most likely due to the specific influence of glasshouse conditions and the unseasonably warm late summer/early autumn of 1999;
- water temperature curves should be slightly delayed. This criterion was not met by the study conditions also due to the glasshouse conditions as well as due to the specific temperature regime of 1999 and 2000.

In summary, the study can be regarded as representative for European climates with respect to temperature dynamics, seasonal variations and temperature extremes.

The main P- and N-nutrients o-phosphate, nitrate, and ammonium were measured weekly. Nitrogen salts are steadily metabolized to N_2 by microorganisms in anaerobic parts of the sediment and eliminated from the systems, a loss which cannot be compensated by atmospheric deposition and natural input from leaves as in outdoor systems. Therefore, N was additionally introduced in the systems as $NaNO_3$ (15 mL of a 50 g/L solution = 10% of the initial nitrate concentration) starting in November 1999, when the nitrate values were considerably below the initial values, and added weekly as long as necessary to maintain the initial concentrations (referring to the practice of TCM Brock, pers. communication). By this measure, the values were kept between 2 and 8 mg/L throughout the study. Phosphate was not regulated. The values varied widely between microcosms and sampling dates. They ranged from 0.1 to 1 mg/L during the treatment period and from 0.1 to 2 mg/L in autumn 1999. During the winter, the variability was lower (0.21 to 0.88 mg/L). In spring and summer 2000, algal blooms and breakdowns caused variability from 0.1 to 1.0 mg/L and 0.9 to 2.2 mg/L. Ammonium was generally below 0.1 mg/L. DOC values were between 6 and 10 mg/L during the treatment period and between 10 and 16 mg/L at the end of the study. During the period of rapid increase in algal populations in early summer, DOC values reached 25 mg/L.

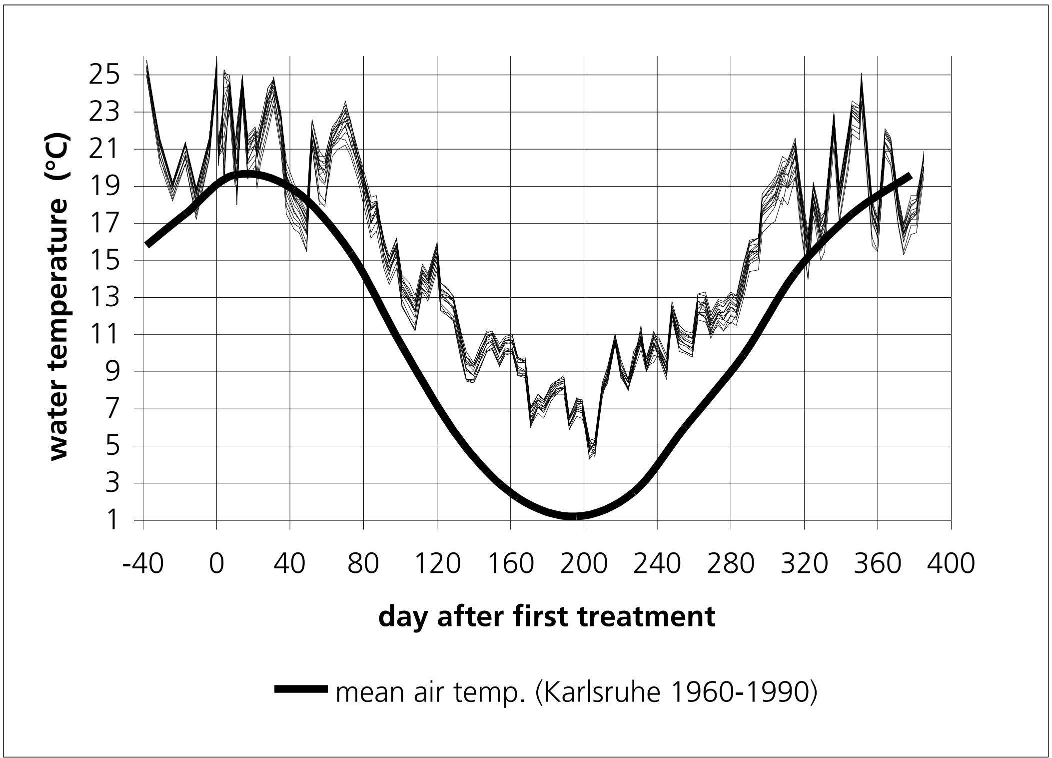


Figure 25: Water temperatures of the 16 microcosms throughout the study, compared with a mean German air temperature regime

Fate results. The nominal concentrations were confirmed by the results of unfiltered microcosm water samples. The concentration of copper in filtered samples, excluding particle bound copper and adsorptive colloids, tended to decrease with further applications (Figure 26). The decrease of copper in filtered samples of surface water is considered an indirect effect of copper itself by reducing populations of predating zooplankton and activities of competing macrophytes, resulting in algal blooms. The significant increase in algae resulted in an increase in surface area and algal exudates, both of which bind copper. Four weeks after the last application the water concentrations of the 0.45 µm membrane filtered fraction at the highest dose level was below the level of the lowest observed effect concentration. Within that time, the sediment concentrations and pore water concentrations of the upper 5 cm layer at the two highest treatment levels increased steadily to a first peak, being two and five times higher than control values for sediment (30 mg/kg) and pore water (5 mg/kg), respectively. After a decrease in winter, a similar peak was measured in summer of the following year. In macrophytes, copper concentrations increased to values higher than control values by a factor of 100 in the week after the last application. During the next 5 weeks the values decreased by a factor of ten, remaining constant for 250 days until the end of the study with values being 10 times higher than in the control plants. Snails did not accumulate copper considerably. At study termination, in snails of the fate microcosms, copper concentrations were 37 mg/kg compared to 24 mg/kg dry weight in the treated and untreated microcosm, respectively. During application, copper concentrations in the plankton fraction corresponded to the dissolved copper concentrations in water, being 50 –

100 µg/L at the highest treatment level. At study termination, concentrations were 0.1 to 0.5 µg/L in all microcosms and thus lower than the dissolved concentrations in water.

Regarding the mass balance calculation, the copper recovery in the upper sediment layer and the plants was 611 mg after subtraction of the background values as determined in the control microcosm. This corresponds to 114% of the amount applied (536 mg). The copper content of the snails was not considered for the mass balance since the copper amounts were not relevant (too low). The plankton fraction was included in the value for the water phase (non-filtered samples). The lower sediment layer was not considered because it was shown that the copper content of this fraction was not changed during the course of the study.

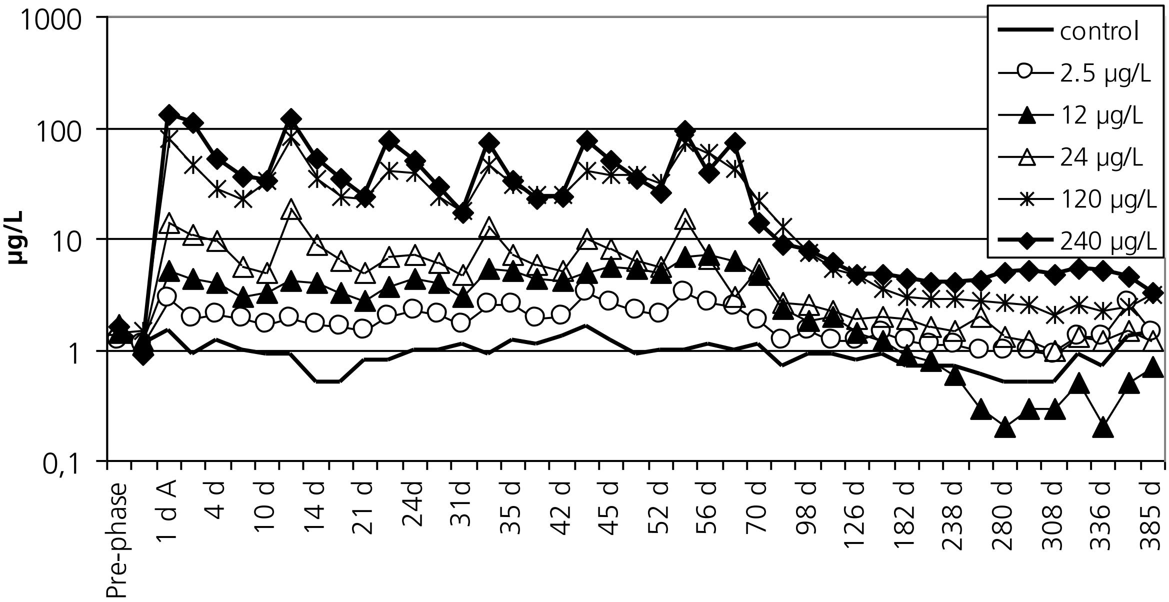


Figure 26: Dissolved copper (0.45 µm-filtered) in the water phase throughout the study

Effect results. Population or biomass decreases were observed for the zooplanktonic Phyllopoda and Rotatoria, the macrophyte, and the phytoplanktonic Cryptophyta. The zooplankton community was shown to be the most sensitive community (Figure 27), the surface grazing cladoceran *Chydorus sphaericus* being the most sensitive species in the study, showing temporary effects at 24 µg/L (reduction of abundance for the first month after the first treatment, recovery within the treatment period). This was also reflected by a single NOEC of the multivariate zooplankton community analysis (Table 12), whereas the other cladocerans were not affected at that concentration. *Chydorus sphaericus* is mostly living in contact to periphyton. The very high copper concentrations in macrophytes at the end of the application period were most probably due to colloidal material adsorbed to the surfaces and might be taken up by *Chydorus* via ingestion. The rotifers (e.g. *Keratella quadrata* or *Cephalodella*) were reduced only at the highest treatment (but without full recovery) and partly profited from the decreased competition by the cladocerans. A total recovery of rotifers might have been prevented by very high abundances of the recovered Phyllopoda which are generally superior competitors compared to the much smaller rotifers. Reduced competition for zooplanktonic species and increased food level can explain the increase of Copepoda (right side of the species weight figure, Figure 27) and, less pronounced, Ostracoda.

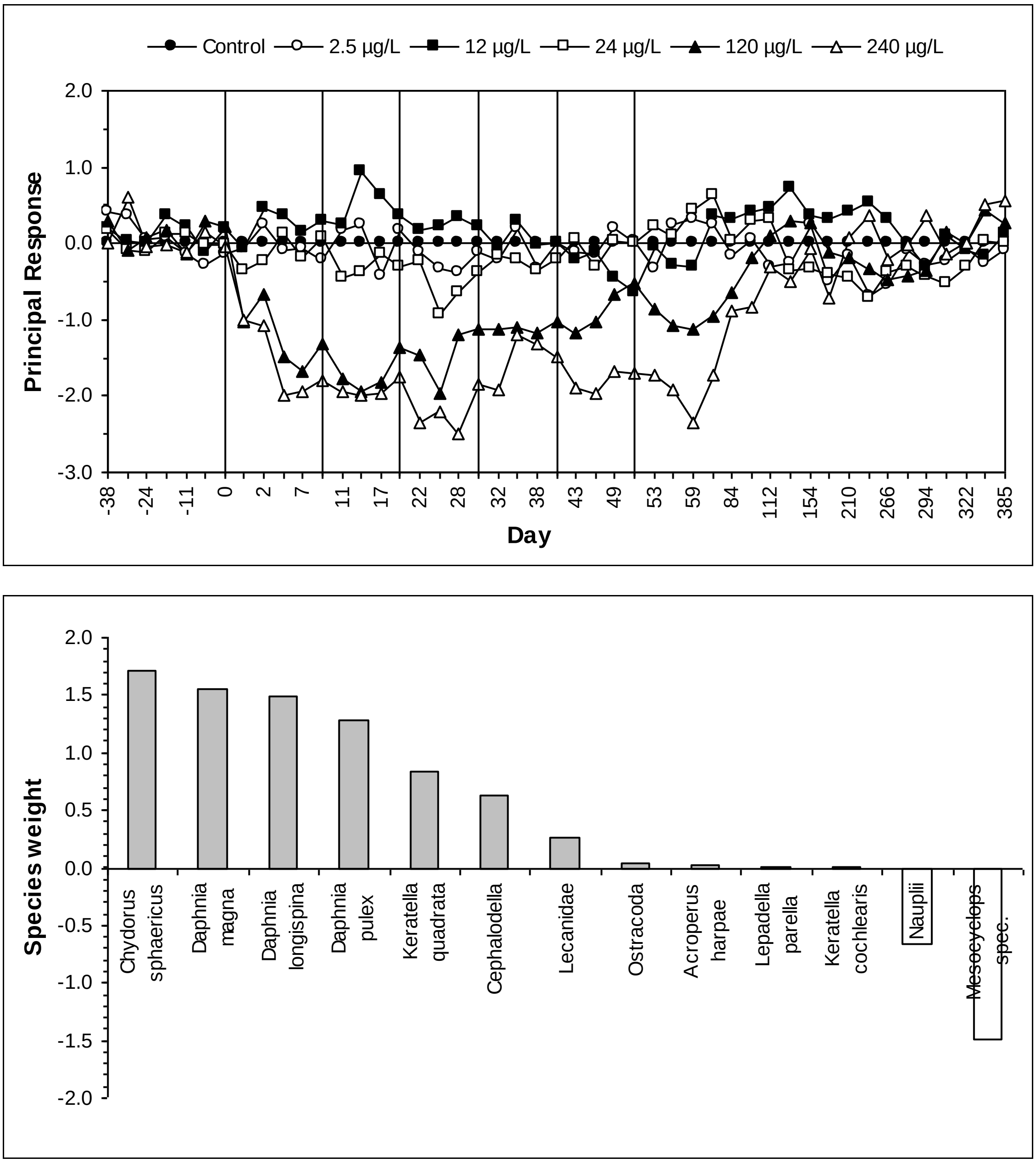


Figure 27: PRCs for the zooplankton (top) and related species weights (bottom)
PRCs indicate a significant part of total variance (p=0.005). Vertical lines indicate days of application.

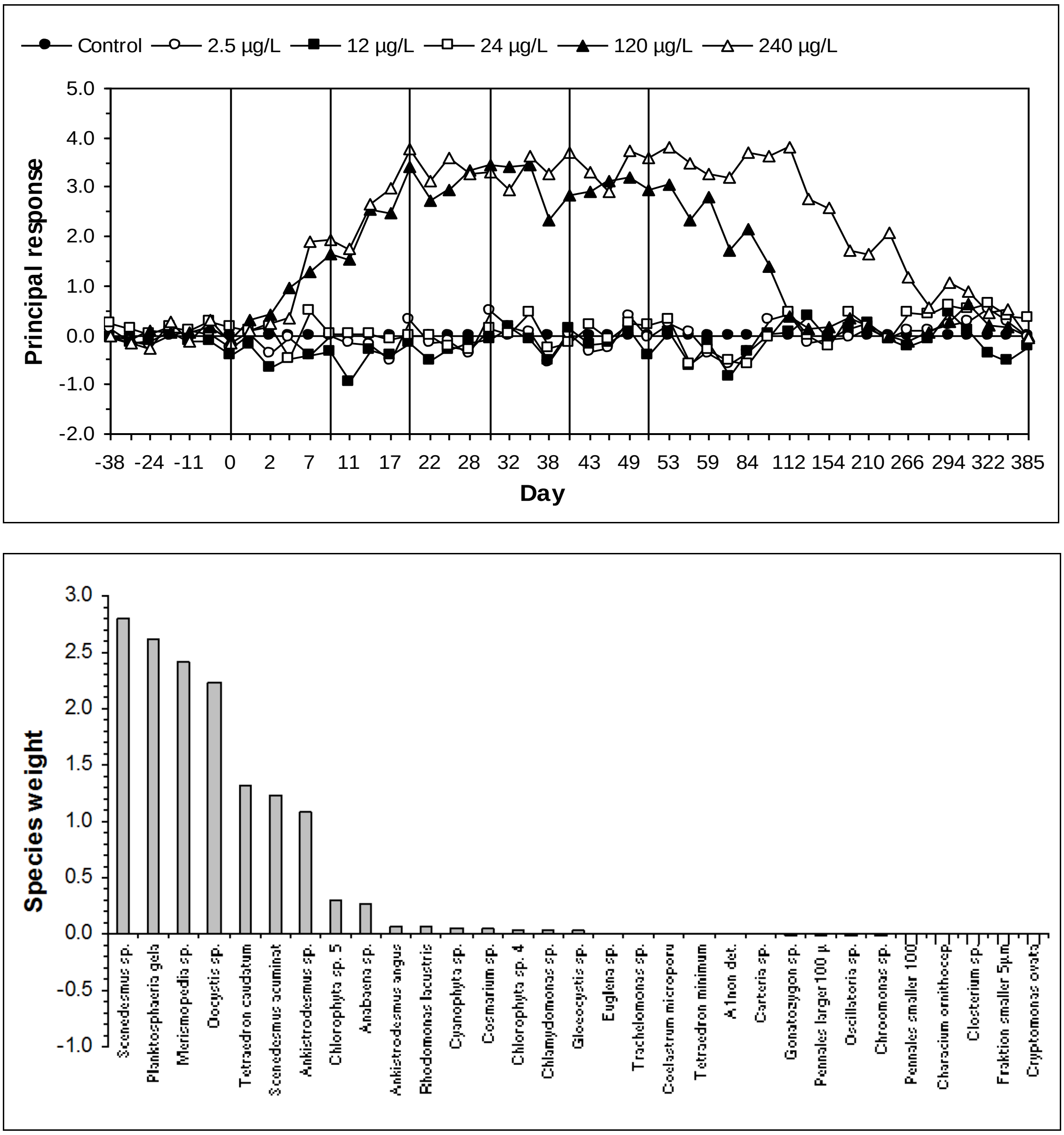


Figure 28: PRCs for the phytoplankton (top) and related species weights (bottom)
PRCs indicate a significant part of total variance (p=0.005). Vertical lines indicate days of application.

The phytoplankton community, dominated by green algae, showed considerably population increase as indirect effect of the decreasing predation pressure by the cladocerans (Figure 28, NOEC 24 µg/L). However, a look at he right side of the species weight figure (Figure 28) revealed that for example the population dynamics of the cryptophyte *Cryptomonas ovata* did not follow that characteristics but showed significant decrease at 120 µg/L (NOEC: 24 µg/L). The high densities of Chlorophyta might have caused further indirect effects due to competition of primary producers: Whether less increase of more r-strategic and less competitive Cryptophyta at the highest concentration was more due to resource competition or toxic effects could not be explained. The effect on the biomass production of the macrophyte *Elodea densa* (NOEC: 24 µg/L), which more clearly was a direct effect, certainly was supported by light and nutrient shortcomings caused by planktonic algal blooms.

Table 12: Comparison of NOEC and NOEAEC (according to Brock et al. 2000) values calculated by using different community evaluation methods

Evaluation method	**NOEC**	**Highest concentration with full recovery**		
		Within the treatment period	**Within 8 weeks after the last treatment (d 112)**	**Until the end of the study**
Zooplankton community				
Number of taxa	24 µg/L	120 µg/L	120 µg/L	240 µg/L
Diversity (Shannon-index)	24 µg/L	240 µg/L	240 µg/L*	240 µg/L
Similarity (Steinhaus-index)	24 µg/L	120 µg/L	240 µg/L*	240 µg/L
Redundancy analysis (PRC) PCA 1st axis sample scores	12 µg/L	24 µg/L	24 µg/L	240 µg/L
Most sensitive taxon (*Chydorus sphaericus*)	12 µg/L	24 µg/L	240 µg/L	240 µg/L
Phytoplankton community				
Number of taxa	24 µg/L#	240 µg/L	240 µg/L**	240 µg/L
Diversity (Shannon-index)	240 µg/L	240 µg/L	240 µg/L	240 µg/L
Similarity (Steinhaus-index)	12 µg/L	24 µg/L	120 µg/L	240 µg/L
Redundancy analysis (PRC) PCA 1st axis sample scores	24 µg/L#	24 µg/L#	120 µg/L#	240 µg/L
Most sensitive taxon (*Cryptomonas ovata)*	24 µg/L	24 µg/L	120 µg/L**	120 µg/L
Macroinvertebrates				
Number of taxa	24 µg/L	240 µg/L	240 µg/L	240 µg/L
Diversity (Shannon-index)	120 µg/L	240 µg/L	240 µg/L	240 µg/L
Similarity (Steinhaus-index)	240 µg/L	240 µg/L	240 µg/L	240 µg/L
Redundancy analysis (PRC) PCA 1st axis sample scores	n.s.	n.s.	n.s.	n.s.
Most sensitive taxon	Not reliable due to very low numbers of individuals/taxon			
Macrophyte production				
Dissolved oxygen concentration	24 µg/L	240 µg/L	240 µg/L*	240 µg/L
pH	12 µg/L	120 µg/L	120 µg/L	240 µg/L
Biomass of *Elodea densa*	24 µg/L	24 µg/L	120 µg/L	120 µg/L
Most sensitive value	12 µg/L	24 µg/L	120 µg/L	120 µg/L

Note: In all cases hypothesis testing was performed by using Williams' test
* effect again at 240 µg/L between days 112 and 294; no effect at 120 µg/L
** increase again at 240 µg/L between days 182 and 266
n.s.: not significant
increase

Primary production, as the basic functional parameter of lentic systems, was dominated by the macrophyte, which was demonstrated by the effects on water parameters. In spite of algal blooms, oxygen levels and pH-values decreased at the two highest dose levels indicating a net decrease of photosynthetic action and primary production. Even at 24 µg/L, primary production (measured as decrease of pH of the water column due to reduced consumption of CO_2) was affected during the treatment period, corresponding to the LOEC of the most sensitive species. Thus, macrophyte performance was one of the most sensitive endpoints when looking at short-term effects. With respect to long-term functional effects, the microbial degradation of organic material (macrophytes, plankton) caused further oxygen depletion and increased DOC and ammonium concentrations at 120 and 240 µg/L.

Due to low numbers of macroinvertebrates and high variability between replicates, statistical differences between the copper treatments and the controls could not be found in the sediment organisms, neither in the emerging insects. However, the numbers of Tubificidae tended to increase, especially relative to insect larvae, at 240 µg/L. Odonata were able to emerge at least up to 24 µg/L. Since low densities and high variability of sediment organisms were expected, an in-situ bioassay with a known number of introduced individuals should guarantee clear results. *Chironomus riparius* was chosen as test organism, as it is a well-known sediment dwelling species, which is used within the pesticides notification process whenever a risk for sediment organisms is identified. It is a species with a short generation time and thus suited for being used in a bioassay at defined short periods with different exposure situations. However, during the study some unexpected problems occurred when performing the in-situ bioassays: The first in-situ bioassay was prepared by introducing Plexiglas tubes in the sediment of the effect microcosms 8 days after the last application. The tubes were closed by a net tent above the water surface serving as emerging trap. Two weeks later, 30 first instar larvae were introduced in each tube and fed with TetraMin fish food (quantity comparable to laboratory tests). Temperature was above 20 °C up to this date, but declined rapidly to below 13 °C during the following four weeks. The chironomids were not able to develop fast enough to emerge at sufficient temperatures at all dose levels except the highest one. The tubes were left in the systems until late spring of the following year, when temperatures were again above 17 °C, but the introduced larvae, which did not emerge in autumn, had obviously not been able to overwinter in the systems. In spite of these shortcomings, a clear result could be obtained: The copper concentrations even at the highest dose level were not reducing emergence rates of chironomids. When regarding the test conditions, 33% to 43% emergence at 240 µg/L can be evaluated as good rates, which may have been due to the enrichment of the sediment, by organic material from dead macrophytes, filamentous algae and plankton. The enhanced food level obviously may have provided the larvae with a significantly better diet than the introduced fish food, resulting in emergence in time before the hibernation period. The cold temperature prolonged emergence times to 33 to 34 days compared to three weeks in laboratory experiments at 20 °C. The facts of very close emergence times and very similar emergence rates and sex ratios in the two replicates at 240 µg/L support the hypothesis of indirect (enhancing) effects compared to the control systems. Because of these experiences, the second in-situ bioassay, which was planned in late autumn, was left out. The last in-situ bioassay was prepared by introducing the tubes on 23.05.2000 and started by introducing 30 first instar larvae in each tube two weeks later. Temperature was between 16 and 24 °C until the end of the study. The results, however, indicate that food supply again seemed to be the factor dominating emergence success. This time, also at the two lowest dose levels some individuals emerged. Emergence started on day 21 of the bioassay, being comparable to standard test results. However, emergence times were extended to 41 days, and again, no control chironomids emerged and the rate was clearly the highest at the highest dose level of 240 µg/L. At that dose level, the mean of the replicates was similar to that of the first bioassay, but the variance between the replicates was very high. The high rate of 67% was achieved in the microcosm with the lowest oxygen saturation during the time of the bioassay, indicating microbial degradation of organic material left from the effects of the year before, resulting in the presence of food resources for the chironomids. The conclusion of the in-situ bioassays is the necessity of improving additional food supply and taking care of the temperature regime at bioassay performance. With respect to the objective, no direct effect on the emergence at dose levels up to 240 µg/L could be observed. The same observations were made with small

Brachycera, also emerging in the bioassay tubes. Exposure via sediment seems to be of less importance than exposure via water.

For the treatment period, the structural NOECs (single species as well as the community NOEC) of the most sensitive community and the functional NOEC is set to 12 µg/L. When looking at recovery times, a clear relationship to treatment levels was observed. All significant effects and trends, which could be observed at 24 µg/L, recovered within the treatment period (most sensitive NOEC: 24 µg/L). All effects observed at 120 µg/L recovered within 8 weeks after the last treatment in autumn. At study termination after one year (start of the next potential treatment), all endpoints showed clear tendencies of recovery even at 240 µg/L. However, recovery was not completed for macrophyte (*Elodea densa*) biomass production, because it is much less dynamic than the planktonic communities are; the production rate was comparable to those of the other systems in spring and early summer following the year of applications. Thus, the NOEAEC is 120 µg/L.

Extrapolation to other products. As other copper-based fungicides act through the same mechanism by setting free Cu^{2+} ions, the results of the study can be extrapolated to different copper salts (hydroxides, oxychloride, sulphates, oxides) and different formulations by conducting appropriate bridging studies. These studies should represent comparable conditions in terms of water and sediment quality, the most sensitive organism group and the same application pattern. For this, special 21-day reproduction tests were performed with *Daphnia magna*, which was shown to be one of the most sensitive species in the microcosm study. The test vessels contained a 2 cm layer of the same sediment from the same pond used as sediment source in the microcosm study. The tested nominal concentrations were 2.5, 12, 24, 120, and 240 µg/L; the medium was renewed on the tenth day to generate exposure conditions comparable to the microcosm study. After filling the prepared test dispersions in the 250 mL test vessels already containing sediment, cylinders made of thin Plexiglas and a fine gauze net at the bottom were placed in the test vessels in order to be able to take out the newborn daphnids.

Measured initial concentrations of total copper were within +/- 20% of nominal concentrations at least for those treatment levels that affected the test animals. In all tests, survival and offspring numbers fulfilled the validity criteria of the usual Daphnia reproduction test without sediment according to OECD TG 211.

Survival was identified as the most sensitive endpoint. Generally, the tested oxychloride and oxide products were less toxic than the other products with NOECs for survival of 120 µg/L compared to 24 µg/L. Reproduction was not significantly reduced when observed in animals surviving over the whole test period of 21 days. Animals, which survived at 240 µg/L, showed no reduction of reproduction. The intrinsic rate of increase, which combines age specific survival and fecundity in order to assess possible effects on population dynamics, delivered equal NOECs to the ones calculated for survival.

Because the copper hydroxide product tested in the microcosm study showed similar or higher toxicity compared to the other copper salts, and because sensitivity of *D. magna* in the reproduction test and Daphnia populations in the microcosm study were comparable, the modified reproduction test is regarded as representative for the microcosm study. A risk assessment based on the microcosm study can be extrapolated to the other tested products.

2.4.4.4 Extended fish studies

Whereas higher tier zooplanktonic community tests including recovery by repopulation are accepted, the situation for fish is different. Due to the long generation time, recovery in due time from effects on survival is not possible, from effects on growth it is difficult. Prolonged toxicity tests as OECD TG 204 or 215 are commonly performed under flow-through and sometimes under semi-static conditions. In the case of test substances, which are rapidly dissipating from the water phase, the common laboratory exposure represents an unrealistic worst-case situation, resulting in unrealistic sensitivity to long-term effects.

For example, acaricides acting as electron transport chain inhibitors represent a ubiquitous mode of action with comparable effect thresholds in many heterotrophic species. These active substances are designed to be adsorptive and are often very lipophilic. Thus, the dissipation from the water column is fast and the exposure to peak concentrations short. The species at highest risk are those with high activity and water exchange rates at respiratory membranes, such as fish and Daphnia (see chapter 2.2.3). For refinement of the risk to fish populations, the exposure has to be focused on. However, it is not clear whether a short-term exposure to a lipophilic compound may also result in accumulation and cause long-term effects. Extended laboratory studies with a more realistic exposure reduce the uncertainty of extrapolation from laboratory to the field. Additionally, they can be used to enhance the NOEC used in regulatory evaluations. As they are still single species studies, they are not suited to reduce the uncertainty concerning sensitive species and most sensitive endpoints and thus to reduce the acceptable TER value.

We used our aquatic microcosm facility for a test method specifically designed for the refinement of the chronic fish NOEC for pesticides (ref. 83)- 85)). For this, the species found to be most sensitive in the lower tier tests has to be included, as it is hardly possible to replace an existing NOEC by a higher one of a more realistic study, if it was generated using a different species. A realistic worst-case exposure scenario for aquatic organisms is the overspray of a lentic shallow drainage ditch adjacent to the crop, which can be simulated by lentic meso- or microcosms including sediment and macrophytes. Unfortunately, in most cases the species needed is rainbow trout (*Oncorhynchus mykiss*), which is a fish species of lotic waters, needing comparably low water temperature and, if tested in lentic systems, sufficient water to move in. This excludes outdoor mesocosms due to temperature needs as well as laboratory systems due to space. Our microcosm facility provides sufficient water for a static exposure of rainbow trout and the possibility of managing the water temperature. As a result, the unrealistic exposure of testing a lotic fish in lentic systems is the most appropriate way to find a compromise between realism and worst-case exposure, which is accepted by industry and authorities.

Testing juvenile growth of rainbow trout in a semi-realistic microcosm needs additional feeding and access to the fish for weight measurements. In a complex system, this is only feasible when encaging fish. For this, we used covered stainless steel net cages (40 x 40 x 30 cm l x d x h) with 5 mm mesh width and a fixed cover. Another complication is the habituation of trout to feeding, resulting in active predatory behavior at the water surface, whenever a person approaches the microcosm. To keep the habituated fish off the water surface when applying the test substance, the closed cages were placed beneath the water surface in the centre of the water body during and shortly after application.

After introduction of water and sediment from natural sources (see chapter 2.4.3) and during the pre-treatment period, the microcosms were connected to one another by a tube system ensuring a slight water exchange between the systems driven by aeration. For this reason, the water was filled up to 10 cm below the brim. Shortly before treatment, the systems were separated and the height of the water body was adjusted to the scheduled size. Within the pre-treatment period, macrophytes (*Elodea* spec.) were introduced into the microcosms. Plants were fixed on a stainless steel wire mesh (mesh width 2 x 2 cm) at the bottom. Density and total biomass were approximately equal in each microcosm (120 plants of about 30 cm length per replicate). The wire mesh was placed on the top of the sediment. Additional macrophytes being introduced with the sediment by chance were eliminated some days before treatment in order to create equal conditions. At least 10 days before treatment, 10 fish (*Oncorhynchus mykiss*, 5 ± 1 cm) were introduced encaged in each microcosm to study the acclimatization process to the microcosm conditions and to give the opportunity to optimize the system and/or replace weak or dead fish in time. The cages were fixed at the centre of the water surface until shortly before treatment to ensure a sufficient oxygen supply. Additional feeding consisted of commercial trout pellets (approximately 5% of the initial fish weight per day).

Shortly before treatment, the aeration of the fish cages, the water exchange between the microcosms, and the additional feeding were stopped. During the application period and the following 48 hours, the fish cages were fixed in the centre of the water body (Figure 29) to avoid direct contact of the fish to overspray during treatment. The fish were surveyed visually through the glass panes. 48 hours after treatment, the cages were fixed at the water surface. A slight aeration and additional feeding were started again after 96 hours.

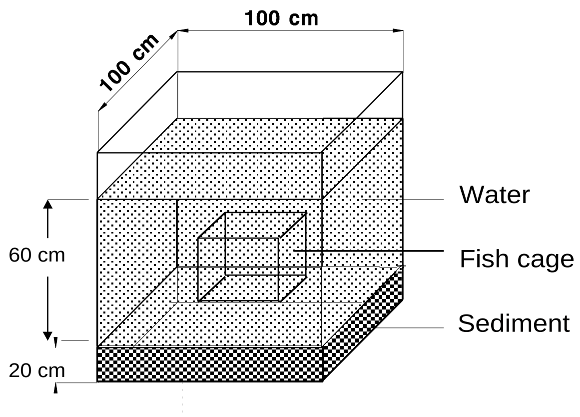


Figure 29: Extended fish study in the IME microcosm facility: Placement of the fish cage during and for 48 h after treatment. Organic carbon content of the sediment was approximately 3.5%.

Depending on the test design triggered by the required statistical evaluation, five to six nominal test concentrations were applied without replication or in two replicates with up to four controls. The systems were treated with the formulated product to achieve a realistic distribution after spray drift exposure. “Nominal” means the theoretical concentration of active substance assuming a complete solution and an equal distribution of the applied amounts in the whole water body. A microcosm at the highest or second highest concentration could be used for fate analyses by applying formulated ^{14}C-labelled active substance[6] as done in the presented example (ref. 83).

The application solutions were applied to the water surface by means of separate hand-sprayers. After application, 20 mL of water were added to each hand-sprayer and after intensive shaking by hand, this solution was applied to the water surface too. Directly before

[6] The formulated ^{14}C-labelled application solution was prepared by reducing an appropriate amount of the delivered solution of ^{14}C-labelled active substance to dryness and dissolving the residue in 500 µL acetone. 10 µL of this solution were used for the determination of ^{14}C-radioactivity by LSC. 459 µl were mixed with 116.3 mg pure formulation and 75 mL water in the hand-sprayer (Biomat 2-1980, NeoLab, D-69123 Heidelberg. The formulated unlabelled active substance was directly mixed with 75 mL of water in separate hand sprayers.

the application, the concentration of the application solutions was determined to verify the loading. For that, aliquot samples were taken from the content of each sprayer.

Location of sampling was randomized within each microcosm. From scales at the frame of the vessels, each sampling point could be defined to ± 1 cm. In the microcosms applied with unlabelled test substance, water samples were taken by means of thin glass tubes that were introduced in the water phase cautiously in order to minimize disturbance. On each sampling date, six times 100 ml of water were sampled from each microcosm at water depths of 10 cm and 30 cm, respectively. The six samples from each depth were pooled after sampling. The samples taken at 3 h and 6 h after application were added with an aliquot of a stock solution containing mercury dichloride ($HgCl_2$) in order to ensure conservation. Afterwards, the samples were stored in a refrigerator at 4°C in the dark. The samples were worked up after 8-9 days for chemical analysis. The pooled samples taken later than 6 h after application were worked up immediately for chemical analysis.

In the microcosm applied with radiolabeled test substance, all water samples were taken and treated as the unlabelled pooled water samples taken later than 6 hours. In addition, sediment samples were taken from the upper 0-10 cm sediment layer in triplicate by means of a boring bar. The triplicate sediment samples were pooled before further analysis. Furthermore, three to five (due to estimated mass) individual macrophytes were harvested according to the time schedule. Suspended particles were separated from the water phase by centrifugation. When further work-up was not possible immediately after sampling, the samples (e.g. plant, sediment and particles) were stored frozen at -20°C. At the end of the study and after the water phase had been removed, sediment samples were taken from the upper layer (0 - 5 cm) and from below (5 - 20 cm). In addition, coating on the glass pane was removed by scrapping. Dead fish were stored frozen at -20°C. From an aliquot of the centrifuged sediment samples and of the plant samples, the percentage of the dry weight was determined to calculate a mass balance.

Concerning all physico-chemical parameters during the pre-treatment period, the systems ran mainly parallel, when the systems were connected, as well as concerning all parameters except oxygen after treatment, when the systems were isolated from each other. Main trends were a slight but steady decrease of conductivity, an increase of oxygen concentrations from approximately 8.5 mg/L to around 10 mg/L, an increase of pH from approximately 8.0 to around 9, and an increase of DOC by a factor of 2 during the study. All these trends were caused by the growth of the introduced macrophytes and the resulting consumption of nutrients (including CO_2) and production of oxygen and organic carbon, in combination with the introduction of carbon sources by fish feed.

The mean light intensities varied from about 30 000 lx at the beginning of the pre-treatment period to 20 000 – 25 000 lx during the post-treatment period and were influenced by the outdoor situation, where the light intensity was about 40% at the beginning and only 5% at the end of the fish study due to the season. The systems differed by a factor of at maximum 1.8. The reduction of the light intensities by the distance from the lamp was approximately linear with a lower decrease in water than in air: The water surface is the most important barrier for light energy because of reflection. The longer wavelengths that passed the surface were not absorbed by water as effectively as the shorter ones were by air.

Fate results. In the dissolved water phase of the microcosm treated with 10 µg/L radiolabeled test substance, high concentrations were measured after 3 hours in both water depths, especially in 30 cm depth, reaching more than 100% of the nominal concentration. The decrease of the test substance concentrations occurred rapidly during the first day after application, but afterwards the test substance disappeared more slowly. The test substance occurred mainly in the water phase and in the sediment phase of the microcosm. In the plankton phase the test substance concentration was more than a factor of two lower than in the dissolved water phase. Significant amounts of metabolites were detected in the dissolved water phase after 7 days.

After application on the water surface, large amounts of the test substance were transported to the upper layer of the sediment phase of the microcosm treated with nominally 10 µg/l radiolabeled test substance. In the sediment phase, the test substance concentration reached the maximum of about 35 µg/kg dry weight after 96 h. The distribution pattern over time was similar in sediment and macrophytes, where it was more pronounced (maximum of approximately 10 mg/kg dry weight). In both materials, the maximum concentration was reached after 4 days and decreased by a factor of 10 and 3 in sediment and macrophytes, respectively, at study termination.

The choice of the radiolabeled microcosm at nominally 10 µg/l was driven by the analytical quantification limits for all matrices (with the exception of water) in the non-labeled microcosms and the assumption, that no fish would survive and fate sampling would not disturb effect observations. The dead fish in the microcosm applied with 10 µg/l ^{14}C labeled test substance were sampled after 48 h. The concentration of radiolabeled material was lower than in plants and sediment.

After the end of the study, after 29 d, the microcosm applied with ^{14}C-labelled test substance was disassembled. The ^{14}C-radioactivity was determined in the water phase, in the centrifuged plankton, in sediment (separated in the 0-5 cm and in the 5-20 cm layer), macrophytes, fish (sampled after 48 h) and periphyton on the glass panes. At the time of disassembling, the microcosm consisted of 520 l of water in the water column (losses due to water sampling and evaporation), 42 kg dry sediment of the upper 0-5 cm layer, 154 kg dry sediment of the 5-20 cm layer and approx. 30 g dry plant material. The area of the vertical glass panes, which was in contact with the water phase, was 24000 cm^2. The main part of the radioactivity was still in the water column (47%). 14% of the applied radioactivity was distributed in the upper 0-5 cm sediment layer and 27% was transported to the deeper 5-20 cm sediment layer. The relative and total amount of the radioactivity in macrophytes, fish and on the glass wall of the microcosm was low (2%). In the water column, the radioactivity was determined with less than 10% in the particle fraction.

Discussion. Due to substance properties, an inhomogeneous distribution in space and time occurred during the first days after treatment. The exposed fish were encaged at a water depth of 30 cm. The peak concentrations measured at this depth occurred around 24 h, corresponding to the cumulative fish mortality findings. The measured peak concentration reached 70% of the nominal concentrations. Considering that the real peak concentration might not have been met by the sampling schedule, there is evidence for the assumption, that the peak concentrations differed not more than 20% from the nominal concentrations.

Conclusions. In the acaricide studies, the active substance dissipated from the water phase in a way, that the most serious effects were caused by the peak concentrations which occurred in the critical water depth (centre of the water column) after 24 hours. The peak concentrations were close to the nominal concentrations. The steep concentration-response relationship resulted in acute LC50 values, which were close to laboratory acute LC50 values. At high test concentrations, fish surviving the peak concentration partly showed clinical symptoms within the first week after treatment (retarded movements or reduced feeding behavior). After one week until the end of the study, no abnormal behavior of any of the surviving fish was observed. The fish weights and lengths at the end of the study as well as the pseudo-specific growth rates did not show any significant differences to control fish. Thus, no sublethal effects could be observed at the end of the study, clearly differing from results of prolonged laboratory tests under flow-through conditions.

2.4.4.5 Macrophyte studies

When assessing the environmental risk of pesticides, tests of unicellular algae represent aquatic macrophytes. For herbicides, also *Lemna* is tested. Due to its growth rate, however, *Lemna* often reacts more similar to algae than to macrophytes regarding sensitivity as well as recovery. Thus, frequently *Lemna* is more sensitive than assumed for other macrophytes and the uncertainty concerning the hazard potential for more slowly growing aquatic macrophytes has to be reduced by appropriate studies. These may be studies on species sensitivity distribution of growth effects. At the same time it might be necessary to demonstrate recovery in due time after effects on growth, especially for *Lemna*.

For simultaneously addressing the mentioned aspects, a multi-species test with aquatic macrophytes was developed. It should 1) contain sufficient representatives of different macrophyte taxa including *Lemna,* 2) provide realistic nutrient and exposure conditions by including natural sediment, and 3) be of sufficient duration for a full recovery.

The test was conducted in 270 L-aquaria containing a layer of 3 cm of natural sediment from mesocosms with growing macrophytes. The test item was an organic growth inhibitor, which was applied once in five concentrations and two replicates, followed by gentle stirring (toxicological approach). Test duration was 56 days. The aquaria received ten apical shoots each of the dicotyle species *Ceratophyllum demersum* and *Myriophyllum spicatum,* of the monocotyle species *Elodea canadensis* und *Lemna minor* (20 fronds), and of *Chara intermedia* as alga being similar to vascular plants. During exposure, the dicotyle species *Potamogeton filiformis* developed from the sediment. A second test with a different objective was conducted with copper sulphate, which was dosed two times weekly to simulate a permanent exposure for 28 days (Dieterich 2006). Besides the introduced species mentioned, the aquaria received a further representative of the genera *Elodea* und *Myriophyllum*, *E. densa* und *M. aquaticum*.

Due to longer study duration, the macrophytes developed more biomass in the study with the growth inhibitor (Figure 30). In both tests, the macrophytes showed no effects at the lowest test concentration. Despite the very different modes of action, considerably similar results were obtained: *Chara* was the most sensitive species, but exhibited a smaller effect at the higher concentrations, when the macrophytes decreased. *Elodea* is the most dominant genus. Its decrease with increasing concentrations means a decrease of competition for the clearly less sensitive *Myriophyllum. Ceratophyllum* is the next dominant taxon. When it is less sensitive compared to *Elodea* (herbicide study), it overcomes this taxon at higher concentrations. When it equally sensitive (Cu study, Figure 31), its decrease is of greater profit for *Myriophyllum*.

The individual test species exhibit characteristic dominance patterns, which interact with effects by chemical stressors or even overrule these effects. *Chara* may serve as an example for the increase of sensitivity at low concentrations by competitive stress as well as for the decrease of sensitivity at higher concentrations by the lack of competition. The limited space of the test system results in higher exposure and stronger competition compared to mesocosm studies and thus in a higher sensitivity.

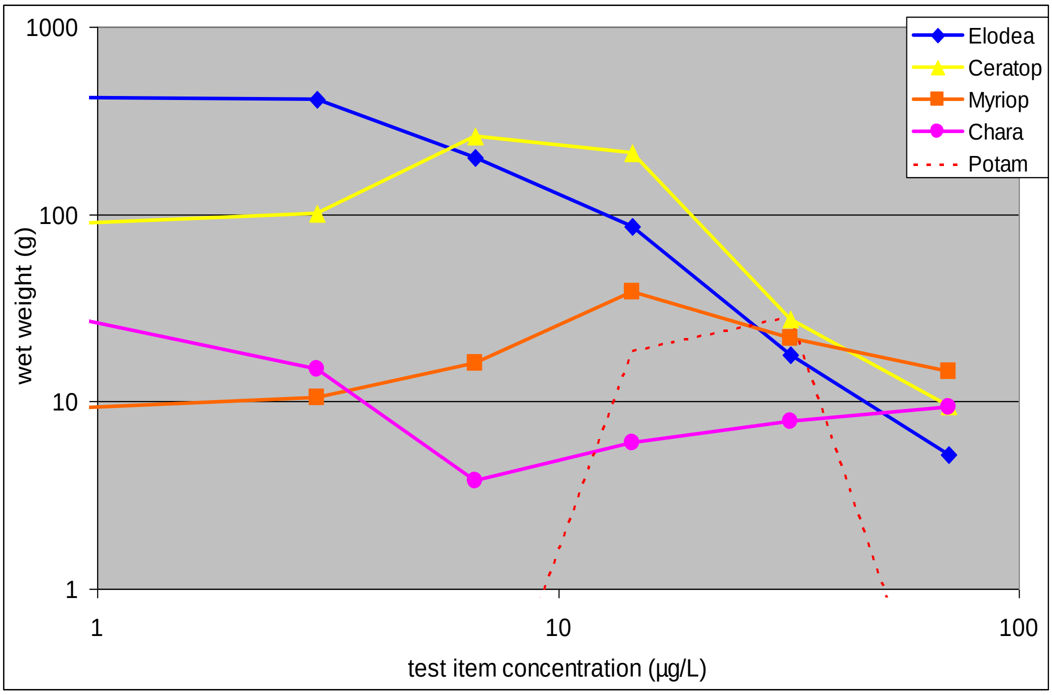


Figure 30: Biomass of macrophyte species after 56 d exposure to a growth inhibitor

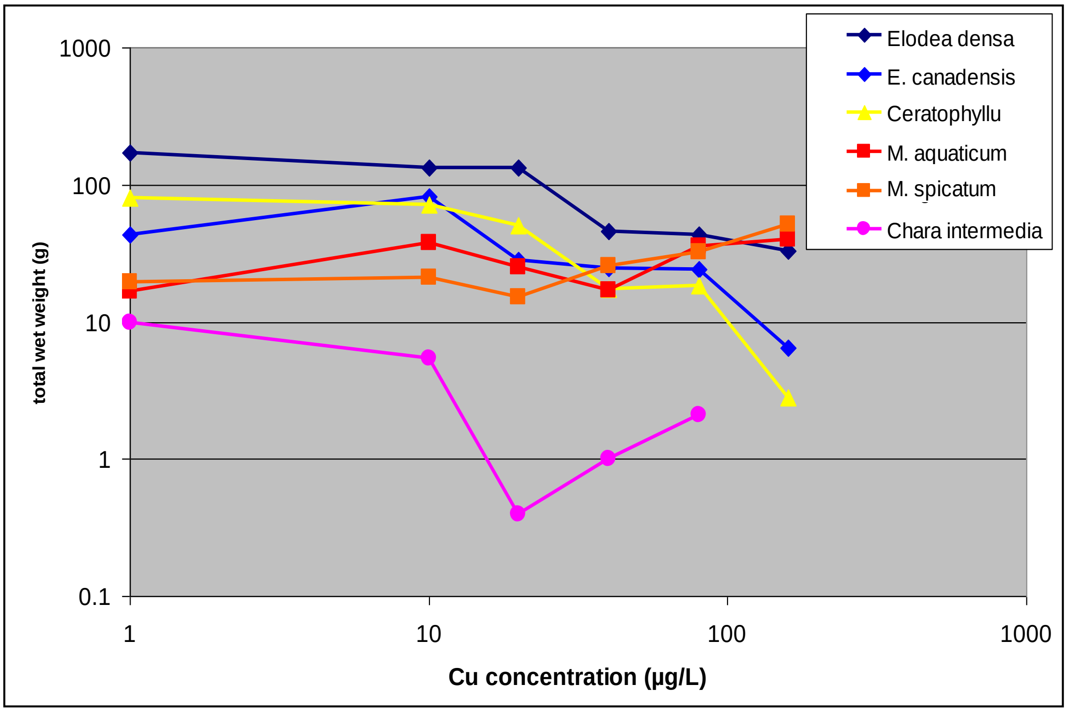


Figure 31: Biomass of macrophyte species after 28 d exposure to Cu sulphate

2.4.4.6 Physical effects of oily liquids

In the last years, paraffin oil products used for crop protection were also subjected to the pesticide registration procedure in Europe. The main ecological side-effect, as demonstrated in standard test, it a physical rather than a chemical effect mainly on daphnids, being trapped in the hydrophobic layer at the water surface. A specific study type was developed at the Fraunhofer IME to quantify the duration of the maintenance of the hydrophobic layer at the water surface as well as the effect on representatives of the most sensitive species and to demonstrate the resulting recovery potential (ref. 71), 76)). The investigated species were the endemic zooplankton community and additionally introduced insect larvae and water bugs (Corixidae). The potential inhibition of three ecophysiological properties was accounted for:

- The emergence at the water surface, mainly by Diptera (*Chaoborus, Chironomus*) and Ephemeroptera (*Baetis*)
- The survival of midge larvae associated with the water surface for breathing (Culicidae, measured as emergence)
- The survival of water bugs (Corixidae) which need to breath at the water surface at regular intervals. The focus was on small species of 0.5 cm length.

For a complete recovery of introduced organisms, the microcosms were fully covered by emergence traps after application. The product was sprayed on the water surface of the indoor microcosms in five replicated concentrations. Four untreated control systems were run in parallel. A further microcosm was used to hold additional organisms for re-introduction into the systems. One microcosm served as fate system for a thorough investigation of the dissipation at the water surface by enclosing a small aliquot of the surface area, introducing defined amounts of cyclohexane including an internal standard, carefully mixing them with the surface layer, removing the cyclohexane and measuring the spectrum of extracted alcanes by GS-MS/MS. As the thickness of the layer is dependent on the water volume beneath, the water column was reduced to 55 cm to approach the regulative standard scenario of 30 cm, but also to ensure sufficient volume for macrophyte growth and oxygen production. This was necessary due to the reduced photosynthetic activity caused by light reflection of the emergence traps.

The water surface was daily observed for dead bugs trapped on the oily surface layer after penetrating it for breathing. Emergence was quantified and the zooplankton community was sampled three times during the first week after exposure and afterwards once weekly without disturbance of the surface layer (funnels fixed at the water surface before application). As soon as the surface layer had disappeared, organisms found impacted were re-introduced. If no further mortality was observed compared to controls, this was regarded as full recovery. This seemed to be justified for organisms that are able to fly and that are characterized by not synchronized reproduction cycles and / or several generations a year (midges and water bugs).

The results indicated that zooplankton, the process of emergence and photosynthesis were not affected by the oil film, whereas survival of Culicidae and Corixidae was reduced in a dose-dependent manner. The highly precise mortality data of Corixidae could be evaluated by probit analysis (Figure 32). The hydrophobic surface layer dissipated in a concentration-dependent manner within a week and full recovery potential could be demonstrated in due time.

Two years later, a second study was performed with a similar product, using microcosms at a water depth of 75 cm. Related to water concentrations, the results for the most sensitive organisms (Corixidae) differed, but related to applied doses per water surface area, they were nearly identical (Figure 32, LD50: 190 mg/m²), underpinning the character of the effect as a physical one.

Consequences for the evaluation of physical effects. Beside the use in the registration of plant protection products, the studies have value for the comparative evaluation of chemically toxic and physical effects in general. When compared to the needed doses to achieve effect concentrations of dissolved substances, the effect doses for the observed physical effects are considerably low. When homogeneously dosing a water body with 100 m² surface area and a water depth of 50 cm with the LD50, an amount of 19 g are needed at the water surface to kill 50% of the most sensitive species. Dosing 19 g of a soluble substance homogeneously to the same water body will result in a concentration of 0.38 mg/L. The data and the comparison support the reported physical effects of non-toxic native oils and fats presented in a review paper prepared for the German UBA (ref. 19). The review paper and the findings were taken as arguments for the 2007 decision of the German commission for the evaluation of substances hazardous for water (KBwS) to include physical effects in the evaluation. In consequence, native oils are no longer classified as WGK 0 (non-hazardous) and thus have to be evaluated in more depth.

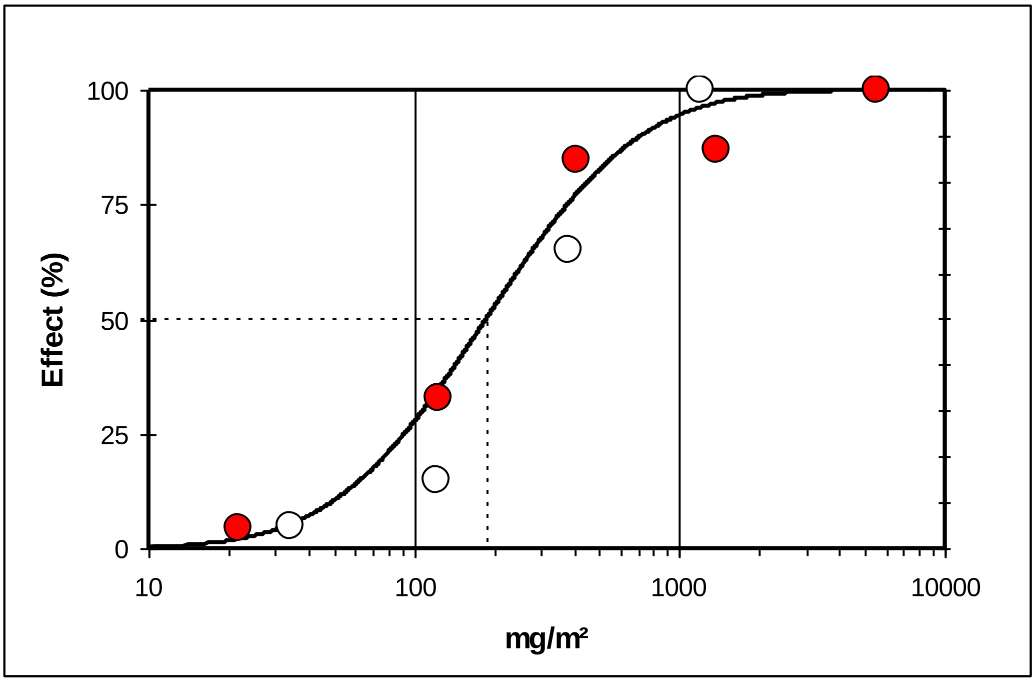


Figure 32: Dose-response relationship of oily surface layers affecting survival of small water bugs (Corixidae) as most sensitive endpoint
Combined statistical evaluation of two independent studies with different water depths, related to applied amounts of paraffin oil products per area. Filled circles: study one; open circles: study two

2.4.4.7 Focused extension of tested scenarios

When evaluating mesocosm studies submitted for pesticide registration, the absolute NOEC or the NOEAEC (including recovery) are divided by the PEC to achieve a TER. The acceptable value of this TER can range between 1 and 10, depending on the quality of the study, i.e., whether

1. the potentially most sensitive organisms are represented,
2. a realistic worst case exposure is represented,
3. the design and performance result in statistical power sufficient to detect relevant effects,
4. the duration is sufficient to demonstrate recovery of affected populations.

If all mentioned requirements are fulfilled by a study, the acceptable TER value including the NOEAEC depends on the data set and additional evidence available to support the study results. For example, the acceptable TER value derived from a single study without additional information about species sensitivities or fate under different environmental conditions is mostly between two and four and is sometimes not appropriate for the registration of specific uses. Additional mesocosm studies representing different conditions are able to reduce the acceptable TER value further. However, they are expensive, time-consuming and often ineffective, as with the knowledge of the first study it can be sufficient to focus the investigations on the most sensitive groups of organisms. Moreover, additional full mesocosm studies can be inappropriate, as specific worst-case study conditions can be of interest, which cannot be realized in a full outdoor mesocosm study.

An outdoor mesocosm study with a fungicide was evaluated as being of high quality. However, as only one environmental scenario was tested, the acceptable TER value derived by the responsible authority was too high to register specific uses. As the most sensitive organisms were rotifers, copepods and snails, which all can be easily investigated in indoor systems, we performed a specific type of study, aiming at a focused extension of the outdoor study by additional scenarios (ref. 79). The main objective of the study was to verify the NOEC and NOEAEC of the outdoor study for further water and sediment qualities with associated communities. Additionally, macrophytes were included which had not been in the outdoor study. To meet the requirements for a high quality study, the following design considerations and measures were made:

1. It should be ensured that rotifer, copepod and gastropod species are sufficiently represented. Additional snail species were introduced and observed not only for abundance, but also for reproduction (egg clutches and juveniles on glass panes).
2. A series of weekly applications represented the worst-case application pattern. As the test substance rapidly dissipates, especially at high pH, the pH was reduced shortly before each application by introducing CO_2. By this, the high pH arising from the photosynthetic activity of the macrophytes was limited for the day of application.
3. Two different sediment qualities including the associated aquatic communities were tested: oligotrophic sandy sediment from the Senne region near Paderborn (TOC of 0.8%) and eutrophic silty sediment from the Vogelsberg region in Hessen (TOC of approximately 5%). Each sediment was tested at the regulatory accepted NOEAEC in four replicates with four untreated control systems. To ensure sufficient numbers of rotifers especially in the oligotrophic system, the sample volume was enhanced to 20 L.
4. The planktonic communities showed fast recovery after slight effects. The snail populations developed equally in the controls and exposed microcosms. However, to ensure the evaluation of potential long-term effects on reproduction including recovery, the observation time was extended for two weeks.

The results could verify effects and recovery of zooplankton populations (NOEAEC) for scenarios differing from the outdoor study by the inclusion of macrophytes, and by testing sediments representing a wide range of nutrient concentrations and particle sizes.

2.4.5 Continuous exposure: Water quality objective setting for copper

When setting water quality objectives for specific substances, the procedure is based on aquatic ecotoxicity tests. The derivation should be based on the best available data set, also including higher tier studies, if performed. However, when including mesocosm studies, the exposure has to be considered with respect to the objective of the risk assessment. For pesticide risk assessment, the initial edge-of the field peak exposure is focused on. Fast dissipation and population recovery can clearly reduce risk, which is the main scope of mesocosm studies.

Water quality standards are used for comparison with measured concentrations in monitoring programs. For pesticides, there may be standards for maximum concentrations and for annual percentiles of regular measurements, which may find concentrations caused by non-agricultural use or by production, storage and transport defects or accidents. The sampling dates are usually not related to application patterns. Especially for rapidly dissipating substances, the probability of finding concentrations in the order of magnitude of the real peak concentration is extremely low. Consequently, it may be justified to use effect concentrations of mesocosm studies based on initial concentrations for the derivation of absolute maximum standards, but definitely not for setting standards for yearly percentiles. For this purpose, either data should be related to time-weighted average concentrations, or they should be generated in studies with continuous exposure. Such a study was performed sponsored by the International Copper Association, represented by the European Copper Institute, to derive a safe concentration for water bodies of copper (ref. 27), 54)). The study was carried out in collaboration with the WRc-NSF, UK, being responsible for the copper speciation measurements.

Copper enters water bodies in large quantities from anthropogenic sources and, unlike organic contaminants, does not decompose. It is highly toxic to aquatic organisms. However, a small amount of the element is essential to all forms of life because copper functions as a co-factor for a number of enzymes and transport proteins. Depending on the geological sphere of influence, substantial background levels of copper may be present, to which the indigenous organisms have adapted. With reference to their biological availability, different metal species occur depending on factors determining water quality (e.g. hardness, pH, DOC).

The objective of the project was to examine the effects of permanent copper-sulphate contamination on an aquatic community. The aim was to verify the ecological relevance of toxicity data obtained from laboratory testing and to extrapolate these to the field situation by testing species from various taxonomic groups and trophic levels comprising all life stages and under realistic exposure conditions.

The direct effects of copper on the abundance of existing species, indirect effects through competition or predation, as well as effects on functional ecosystem parameters like primary production and water quality were determined under reproducible conditions.

2.4.5.1 Procedure

Sixteen microcosms, each 1 m³ in volume, were set into the floor of a greenhouse. They were filled with a layer of sediment (~ 20 cm deep) and a water column (~ 75 cm deep) from a small oligotrophic reference lake near Schmallenberg. The biotic community created by the organisms introduced in this way included phytoplankton, zooplankton and benthic invertebrates. One quarter of the areas was planted with defined numbers and weights of *Elodea densa* and three quarters of the water body was kept free for the plankton communities. Copper concentrations were adjusted to nominal values of 5, 10, 20, 40, 80, and 160 µg/L by three treatments per week and the effects of these treatments were monitored. The application solutions were equilibrated for at least 48 hours before the treatment and then carefully mixed in. The duration of exposure was 110 days. By means of an artificial light and temperature program, contamination was simulated from spring and early summer through to the autumn. To achieve this, the temperature of the systems was first lowered to 6°C and then raised again to 10°C after four weeks. Thereafter, the light and temperature was adjusted in real time.

2.4.5.2 Results

The average adjusted copper concentration measured on the day following treatment, was 70–90% of the nominal concentration. In the 2-3 day interval between treatments, the total copper concentration declined by about 25%. The inorganic copper containing the active form of the free copper ions increased from a measured concentration of about 20 µg dissolved copper/L (Figure 33), which corresponded to the copper complexation capacity of the microcosm water. At the same time, concentrations of free copper ions exceeded 0.01 µg/L.

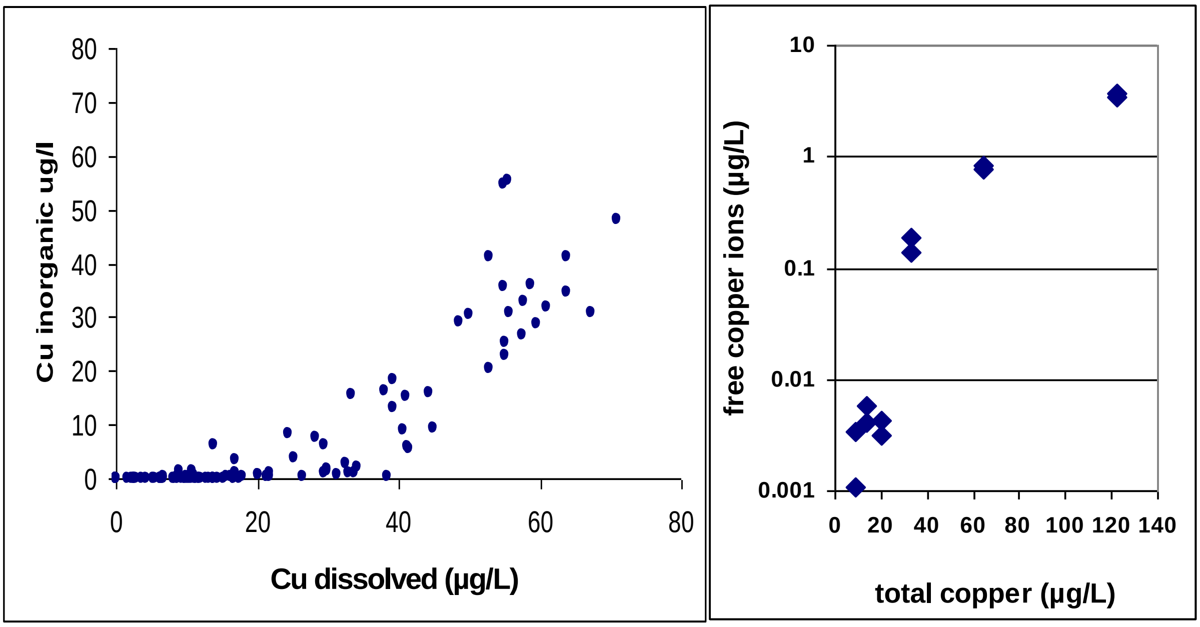


Figure 33: Copper species concentrations in microcosm water samples
Inorganic copper concentrations (containing free copper ions) compared to the concentrations of dissolved copper (left); free copper ions compared to total copper concentrations (right)

The Cladocera, of which *Daphnia longispina* was the most abundant species, were the most sensitive zooplankton species (Figure 34). The NOEC was determined to be 20 µg/L. The Copepoda and Rotatoria species were directly affected only by the highest concentration of copper.

In general, phytoplankton was less sensitive and showed increased densities in concentrations equal to or higher than 20 µg/L due to decreased zooplanktonic predation (Figure 35). The abundance of species of Diatomeae and Cryptophyceae declined temporarily at > 10 µg/L, either through the direct effects of copper or competition through the increased growth of green algae. The current primary production of the systems dominated by the macrophytes, measured by the parameters pH and oxygen content, was already lower at concentrations above 10 µg/L. At the end of the study, the biomass produced by the macrophytes was only significantly reduced at concentrations equal to or higher than 40 µg/L (NOEC: 20 µg/L). For the macroinvertebrates inhabiting the sediment, no effect was observed.

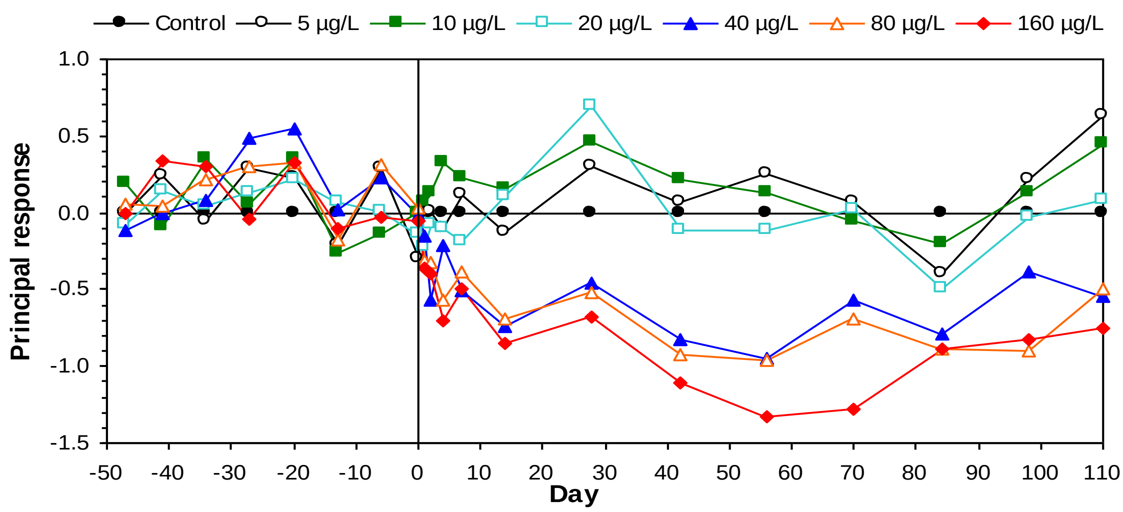


Figure 34: Principal Response Curves for Cu to zooplankton

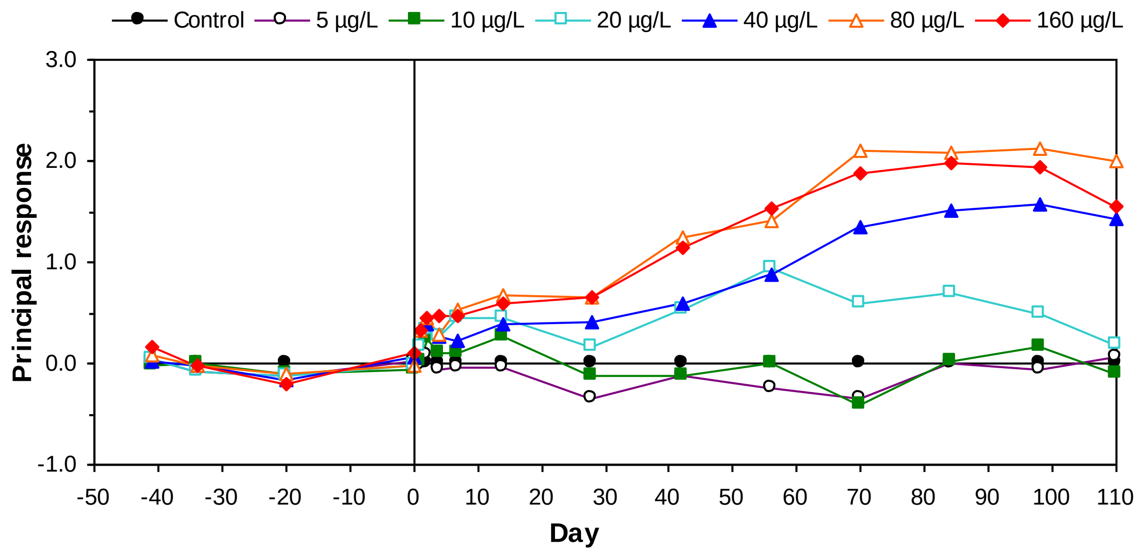


Figure 35: Principal Response Curves for Cu to phytoplankton

2.4.5.3 Discussion of the test conditions

The simulation of seasons by varying the light and temperature produced a good reproduction of successive seasonal events such as algae blooms in spring and autumn. At increasing metabolic rates in spring and early summer, a slight hindrance in the growth of r-strategist populations produced significant effects. In the high summer, when the populations reached their habitat capacities and stabilized at high metabolic rates, the dynamic effect of the season was less predominant. The effects were concealed by variability in the system. In the autumn, falling metabolic rates were reflected in declining rates of growth and compensation potential, which again rendered the effects of copper loads more pronounced. In oligotrophic systems, nutrient concentrations are less variable than in mesotrophic systems, reducing the variability between replicates and optimizing the statistical definition. As the capacity for compensation and the development potential of the populations are lower, the biotic communities appear to be more sensitive.

2.4.5.4 Comparison to results in literature

A literature review of existing aquatic model ecosystem and field test results on copper toxicity (references collected by Patrick van Sprang) was performed to estimate an ecologically acceptable concentration of copper in freshwater ecosystems. Since lentic systems are usually characterized by a higher amount of organic matter in the water column than streams, lentic and lotic systems were assessed separately.

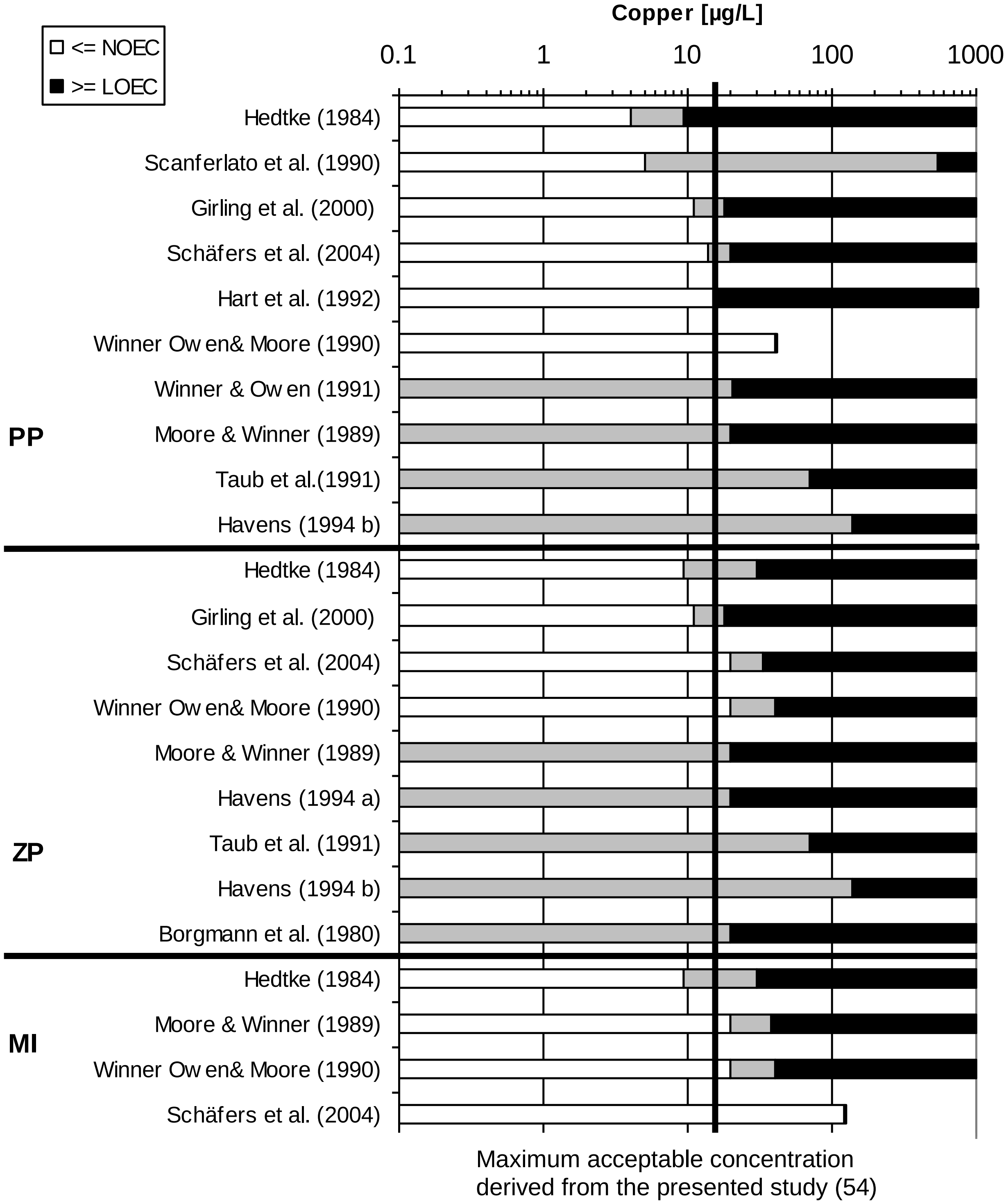


Figure 36: Lowest NOECs and LOECs of published lentic model ecosystem studies with copper
PP = primary producers; ZP = zooplankton; MI = macroinvertebrates

For lentic systems, the lowest observed effects started around 20 µg total Cu/L (Figure 36). The only exception was Hedtke 1984, reporting a LOEC of 9.8 µg/L for a 15 L laboratory microcosm. As this very small planktonic system was exposed under slow flow-through conditions with a DOC of inflow at 0.7-1.8 mg/L, it cannot be regarded representative of natural lentic systems. Winner and co-workers observed effects at 20 µg total Cu/L in spring, but no effects at 40 µg Cu/L in summer due to higher copper complexation capacity (higher phytoplankton density). Macroinvertebrates were generally less sensitive compared to planktonic organisms. Oligotrophic systems seem to be most sensitive. As confirmatory evidence, no effects of copper were observed by Müller et al. (2003) in ditches of the „Altes Land“, characterized by DOC concentrations three times higher than measured in (ref. 54), which is consistent with own considerations (ref. 59).

In summary, for lentic systems with a developed planktonic community, the most sensitive effect threshold concentrations are close to 20 µg/L, but reliability and ecological significance could not always be evaluated.

For lotic systems, i.e., artificial and natural streams, effects were observed at clearly lower concentrations, down to 2 µg total Cu/L. The ecological relevance of effects stated at 2-2.5 µg/L mostly was doubtful (physiological endpoints without effect on production; recovery during exposure; no statistical evaluation). There was no doubt about ecologically meaningful effects starting at 5 µg/L (Figure 37).

The differences between lentic and lotic system sensitivity can be explained when considering copper bioavailability. The low effect concentrations found by Leland and Carter of 2.5 µg total Cu/L were determined for an oligotrophic stream with low copper complexation capacity. The authors calculated the concentration of free copper ions to be 0.012 µg/L. For the clear effect concentration of 5 µg total copper/L, 0.025 µg free copper/L was calculated. These estimations correspond well to the results obtained in the lentic IME study with permanent adverse effects occurring at free copper ion concentrations above 0.01 µg/L.

Clements et al. observed lowest effect concentrations in soft water (53 - 60 mg/L hardness) at 2-3 µg total copper/L. Effects at higher hardness (around 155 mg/L) were smaller: The LC50 for total individuals after 10 days was calculated to be 15 µg/L compared to 6 µg/L in the soft water. Belanger et al. observed clam sensitivity to be less in harder water. The studies demonstrate the importance of water quality for the bioavailability of copper. Unfortunately, neither DOC nor free copper ion concentrations were measured and thus, a comparison with other studies is difficult.

In her PhD work, Hélène Roussell (2005) exposed aquatic communities representative of lentic and lotic habitats in big artificial ditches under very slow flow-through conditions. She stated a NOEC of 5 µg total copper/L and a LOEC of 25 µg/L. The effects were clearer in the upper part of the system near the inlet (copper not complexed), where the system was shallow and more similar to lotic conditions, whereas in the downstream part, the system was dominated by macrophyte and planktonic communities.

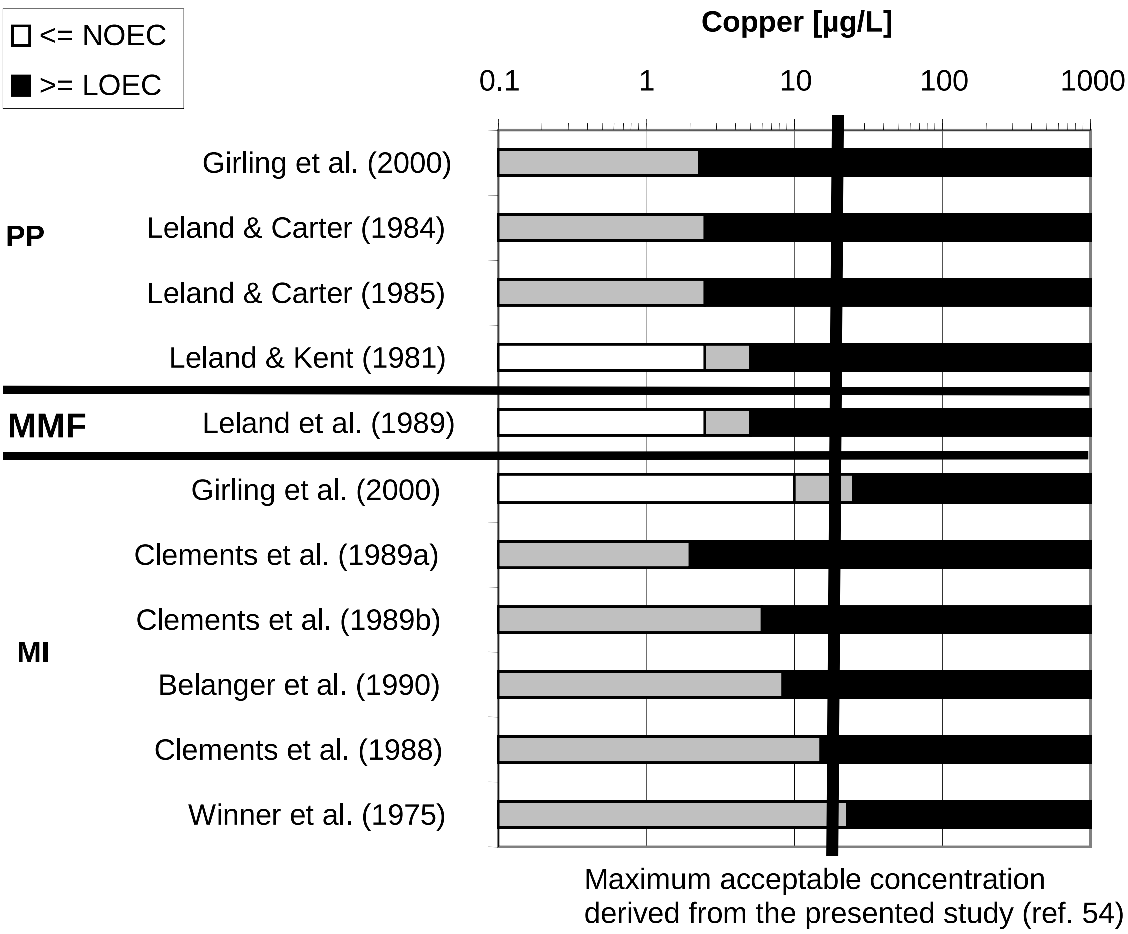


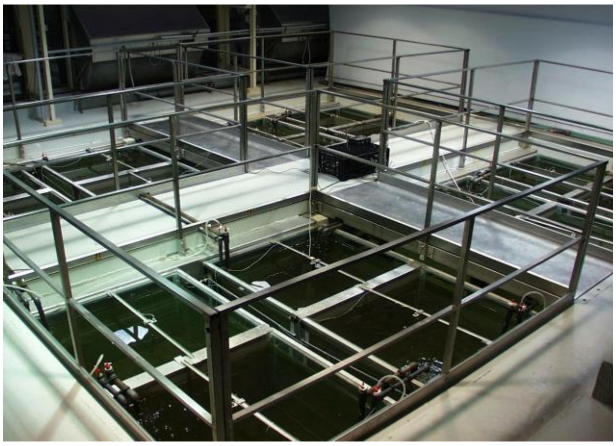
Figure 37: Lowest NOECs and LOECs of published lotic model ecosystem studies with copper
PP = primary producers; MMF = meio- and microfauna; MI = macroinvertebrates

2.4.5.5 Conclusions

With regard to many parameters, the study performed at Fraunhofer IME represents a worst-case scenario for lentic waters: low background values for copper compared to European standards; low water hardness and very low DOC values; the seasons studied; trophic conditions; lack of recolonization; statistical resolution. According to this, added concentrations of copper up to 20 µg/L can be classified as ecologically acceptable. They showed no permanent effect on the structure, and only a very slight effect (with no lasting consequences) on the function of the lentic aquatic ecosystem. This evaluation is confirmed by comparisons with the literature. Due to the presence of planktonic communities, the copper complexation capacity seems to be at least 20 µg/L, which is associated with free copper ion concentrations ≤ 0.01 µg/L (ref. 54).

However, small lotic systems are more sensitive than lentic systems. The driving factor for sensitivity obviously is the availability of free copper ions rather than species sensitivity. The ecologically acceptable concentration depends on water quality, especially the concentration of complexing organic matter. Generally, a concentration of free copper ions of up to 0.01 µg/L can be regarded safe. The effect threshold concentrations of total copper are 15-20 µg/L for plankton-dominated lentic waters and 2-5 µg/L for soft lotic waters with low copper complexation capacity.

2.4.6 Influence of the trophic status (copper study comparison)



The two presented copper studies (ref. 54), 67)) have the following common aspects of design and communities:

- Indoor systems with comparable additional illumination (1 kW metal halid lamps)
- Volume of the microcosms: approx. 1m x 1m x 1m
- Sediment layer of 20 cm
- Design (2 replicates/concentration, 4 controls) and statistics
- Water and included planktonic community of the same source
- Observed endpoints:
  - phyto- and zooplanktonic species
  - benthic macroinvertebrates
  - production of *Elodea densa*
  - water parameters
- Simulated seasons during application (early to late summer)

The microcosms of the two studies were treated several times resulting in a multiple peak exposure period of 60 days in the pesticide registration study and nearly continuous exposure of 110 days in the water quality objective study. Although treated with copper salts characterized by different water solubility (copper hydroxide and copper sulphate), the concentrations of dissolved copper were measured in a comparable range during the period of exposure.

A main difference was the trophic status of the introduced sediment, resulting in different concentrations of DOC and total phosphate in the water columns, whereas no differences in nitrogen parameters were measured (Table 13).

Table 13: Differences in the nutrient status of the two compared copper microcosm studies
The sediment sources were a fishpond and a small pristine lake for the pesticide and water quality study, respectively.

	Pesticide registration study	**Water quality objective study**
Sediment TOC	4.5%	2%
DOC	8 mg/L	4 mg/L
Total phosphate	0.5 mg/L	< 0.1 mg/L
Nitrate	4-5 mg/L	4-5 mg/L
Ammonium	< 0.1 mg/L	< 0.1 mg/L

The effects are summarized in Figure 38. Due to the spacing of concentrations in the pesticide registration study, the comparison is lacking precision. Nevertheless, it can be demonstrated that there is consistency of effects:

- No effect was observed on sediment organisms. It can be concluded that the concentration of bioavailable active copper species in the sediment was not critical.
- The effects plankton and macrophytes were comparable. The formulation of copper hydroxide seems to compensate for solubility.
- The threshold concentration for effects was around 20 µg/L in the two studies. It can be concluded that the copper complexation capacity of oligotrophic lentic systems is able to eliminate bioavailable active copper species up to a total copper concentration of 20 µg/L.

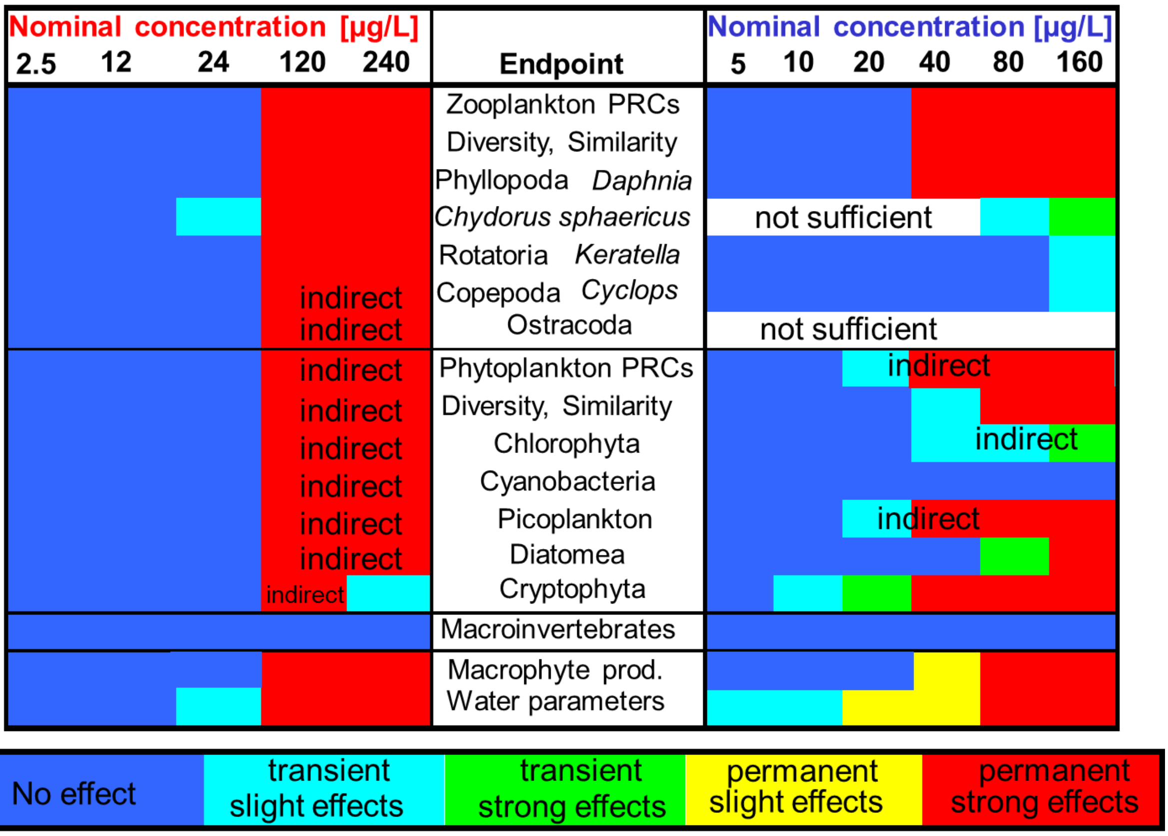


Figure 38: Effect summary of the pesticide registration study (left) and the water quality objective study (right) for the period of exposure
Indirect: Increase of abundance due to lacking predation and / or competition

Some differences could be identified: due to the nutrient status, the plankton densities differed by a factor of five. The effects were more pronounced in the study with higher nutrient concentrations and plankton densities, whereas at lower nutrient concentrations and lower plankton densities the effects seem to be slightly more sensitive. The latter might be caused by the reduced copper complexation capacity as well as by to lower compensatory potential of the communities due to reduced resources and a reduced habitat capacity.

2.4.7 Implications for future work

In the context of water quality objective derivation or REACH, there is value of the indoor microcosm design in confirming no observable realistic community effect levels at no effect levels predicted from species sensitivity distributions. This applies mainly for metal toxicity evaluation. As these NOECs have to be compared with monitoring data, the study has to be performed under permanent rather than peak exposure. Thus, test substances have to re-dosed depending on their dissipation time from the water body, which is more feasible in a well-managed community-level system like our indoor system. A study similar to the copper study (chapter 2.4.5) has already been performed with nickel (to be published in 2012).

In the context of pesticide regulation, the indoor microcosm facility is well suited for specifically designed studies for addressing specific concerns, especially when there is need identified for worst-case scenarios, i.e., southern European climatic conditions, full community level studies at low pH, extended static rainbow trout studies or studies on oily substances. However, its limitations in complexity have to be clearly realized.

The indoor microcosm facility can be used as intercalibration system between mesocosm facilities in different parts of Europe, i.e., for the investigation of global climatic change effects on ecotoxicity. Water and sediment from specific European sources and related communities can be invested at different temperatures at the light regimen needed. This may facilitate a comparison between mesocosm results from different facilities and enabling causal analysis whether the differences are due to species sensitivity, water quality and nutrient status, light regimen or temperature.

The presented results can be used for recommendations for future study conditions, serving specific objectives.

High nutrient levels caused by sediments with a high nutrient status and / or at spring conditions enable high population dynamics. At these conditions, intensive population growth amplifies direct effects on survival, growth and reproduction. At the same time, indirect effects by a lack of predation or competition become dominant and may overrule small direct effects, which might cause problems to show recovery in less complex indoor systems.

Low nutrient levels caused by sediments with a low nutrient status and / or summer conditions limit population dynamics. At these conditions, populations exploit their habitat capacities at a seasonal climax stage. Direct effects on survival, growth and reproduction are less amplified and might be additionally masked by a reduction of intraspecific competition. Indirect effects will be less pronounced. At the same time, direct effects might occur at lower concentrations, as limited nutrition and competitive stress reduce the compensation potential. However, for a detection of these effects, a high statistical power is necessary. Thus, low nutrient levels should better be investigated in managed indoor microcosm studies.

With respect to specific macrophyte studies, a multispecies design in a system with limited resources of space, light and nutrients cannot be recommended, as the competitive interactions between species may overrule substance specific effects. Thus, SSD-approaches (ref. 66) as well as specifically designed outdoor mesocosm studies (ref. 73), 74) with low densities of macrophytes (by excluding *Elodea*) seem to be more appropriate for the assessment of substance-specific risks to macrophytes.

2.5 Realization and communication of risk

The intentional use of ecotoxicologically active substances is a tightrope walk between risk and benefit, the view of which being strongly dependent on the particular interests. By a project of the German UBA (ref. 32) it should be investigated by an interrogation, in how far the users of crop protection products are also sensible for the risks of use. As it was intended to get a cooperative response of farmers about their realization of risk and their creativity of risk mitigation, the interrogation was designed as "Idea competition: Measures for the mitigation of risk of pesticides to ecosystems".

2.5.1 Interviewing farmers on pesticide risk and risk mitigation

During winter 2000/2001 representative German applicants of crop protection products were questioned by a call-centre about their views of risk mitigation. The aim was

- to gain knowledge about the personal attitudes towards risk mitigation measures
- to collect practices of risk mitigation
- to identify differences of attitudes dependent on farm size and cultivation form, on region, and on age of the user

The specially developed questionnaire consisted of three groups of questions:

- Open questions about known practices, starting with a general question being further specified depending on the answer, and ending with special questions about application techniques, application rates, and buffer zones.
- Questions with fixed response choices about the attitude to the actual practice of application regulation (needs, success; responsibility for the protection of ecosystems).
- Open questions about observations of changes in nature and attitudes towards protection targets. The final question aimed at the collection of proposals which had not been called by previous questions, or which had been developed or remembered during the conversation.

About 1600 addresses were chosen containing about 60% arable cropping farms (divided in small farms up to 20 ha, medium farms between 20 and 50 ha and big farms > 50 ha), and wine and orchard cultivating farms with 20% each. It was intended to achieve a representative distribution of farms over the different German Bundesländer. Better accessibility of addresses of medium arable crop farms led to a higher weight of this group compared to small and big farms. The bad availability of valid addresses in the New Bundesländer resulted in a bad representation of these.

Due to the general reserve of the target group and due to the difficult agro-political situation (BSE, increasing threat of food and mouth disease), only 300 successful conversations instead of 1000 expected could be performed. The evaluation of responses focussed on the identification of trends and the collection of practices. Answers to open questions were classified and are presented depending on farm size and cultivation form.

Since hardly any of the small arable crop farmers could be convinced to take part in a conversation, it is only possible to draw conclusions with respect to arable crop farms based on medium and big farms. These and the groups of wine and orchard cultivating farms participated in the interrogation to 21 – 27% of the contacted addresses. Due to their openness, farmers in Baden-Württemberg and Bayern were represented clearly over-proportional and showed the most balanced age structure. The New Bundesländer, Hamburg and Schleswig-Holstein exhibited the lowest response rates and the highest mean age estimates.

As a conclusion from the participation in the interrogation, the results are mainly representative of fulltime-farmers in South- and Southwest-Germany. With respect to Niedersachsen and Nordrhein-Westfalen, the results seem to represent only farmers with special interest in the issue. The northern and eastern parts of Germany are not represented.

Comparable parts of the responding farmers said ‘yes’ and ‘no’ to a possible relationship of the actual practice of crop protection and adverse effects in the environment. If no actual negative effect was stated, this was subjected to the actually notified products being sufficiently caring for beneficial organisms, ready biodegradable, and their applications regulated in a way that correct use causes no adverse ecological effects. If changes in nature were stated, they were mostly addressed to severe changes in selection conditions, resulting in changed species distributions, shifts in dominance and deterioration of diversity. These effects were mainly attributed to herbicides.

With respect to protection needs, different protection concepts can be identified: the major part of responding farmers accepted the obligatory publishing or enlargement of conservation areas. Most of the remaining responses stated that nature as a whole is worthy of being protected and that crop protection should generally account for that, either by careful products, or by giving up chemical crop protection. Especially protection targets which are out of the own observation and evaluation potential (water, especially groundwater), were regarded important, whereas easier observable changes mostly were described as not existent, of minor importance, only temporary or unclear with respect to causality. Several responses identified the lack of long-term observations.

In all responding groups, there was acceptance of the general necessity of application regulations (80% of the wine and orchard-cultivating farmers, 50% of the arable crop farmers). Two thirds of the responding farmers regarded the actual regulations at least as acceptable, with orchard farmers clearly deviated from this trend. All groups ranked actual regulations being „optimal“ below 10%. Generally, co-operation between authorities and farmers and the communication of regulations should be improved, the notification process should be simplified and accelerated, the result should be more practice-oriented and the decisions should be more reliable in the end.

The mentioned possibilities of risk mitigation are of minor importance with respect to conceptual or technical innovations, but give an overview of the diversity of (partly) applied concepts and combination of measures; of the extent of general and special sensibility of the responding farmers to environmental risks; and of the producer’s sentiment towards authorities, market and consumers. Besides many hints to the improvement of application techniques and the variation of application timing, most contributions mentioned variations of application rates: the possibility of an economic use of crop protection products with satisfactory success, at the same time is of ecological use. If such a rectification of economic and ecological interests can be achieved by own management measures, these are favored, whereas the use of innovative application techniques was partly discussed controversially and mostly is outside the own sphere of influence.

Industry was asked to develop selective products caring for beneficial organisms, a broad spectrum of products to provide exchange possibilities if resistances occur, products with long-lasting application successes, and, as the major request, granulated and liquid products. The development of careful pesticides, for example by using natural products, was required. Breeding technology and genetic engineering should be used to develop new varieties, especially resistant to fungi.

Alternatives to chemical crop protection, such as crop rotation, mechanic measures, biological pest control, promotion of beneficial organisms, propagation of new thoughts, represent the most frequent and most diverse proposals. There was a clear request for more research and development of biological pest control products and governmental support.

Politics and markets were perceived as actual hindrance of environmental protection, but also as possible tools for improvements in the future. Distortions of competition within the EU should be reduced. If conditions were equaled in all member states, regulations could be even more strict than today, up to a predominant or exclusive permission of organic farming. Actually, governmental subventions of biological pest control could stimulate farmers. Consumers shoud accept higher costs of sustainable production. Since the government should be interested in healthy citizens, it should support the necessary process of change.

In total, the responsibility for sustainable crop protection mostly was attributed to the product producing industry, with 20 – 30% to the authorities, but with only 10 – 15% to the users themselves. For elder farmers, a general deficit of specific education was stated. The relationship to the authorities seems to be somewhat ambivalent: on one hand authorities are requested to regulate, on the other hand they are criticized because of decisions far from practice and bad information policy. When providing better advice and information, authorities could permit more self-determination and transfer responsibility.

The group of median arable crop farmers represented the biggest part of the participants in the interrogation with the highest estimated age and least dealing with issues of crop protection. The resulting uncertainty concerning some questioned issues was admitted. The part of spontaneous ideas was the lowest compared to the other groups; the part of refused judgments of the regulations as well as the parts of refused estimations of responsibilities and of potential harm of nature by pesticide application were the highest. At the same time, more observations indicating adverse effects were reported compared to the other groups. Improved possibilities of getting advice were required.

The group of big arable crop farmers in many aspects represented the opposite to the median farmers: it is the smallest group, presenting itself as comparably professional in dealing with crop protection, and had the least critical view on potential environmental harm. Education, advice, or information seemed to be no issue in contrast to all other groups. The relative number of responses with respect to buffer zones is by far the highest one: the availability of agricultural area is generally better compared to all other groups. The big arable crop farmers were the only group attributing the main responsibility for the sustainability of crop protection to the authorities. On the other hand, the general acceptance of the necessity of application regulations was comparably low. The possibilities of research and development by the industry were evaluated most positively.

Orchard cultivating farmers suffer most from the high quality pressure by the consumers. They realized a specific necessity of crop protection, resulting in high personal economic loads as well as high potential environmental loads. Among the responding farmers, there was a high part of organic cultivators. The actual practice of application regulations was judged clearly worst compared to the other groups, although the general necessity of regulation is broadly accepted.

The group of the wine cultivating farmers was most communicative and creative. The contentment with the actual situation was clearly better than in the group of the orchard-cultivating farmers.

The interrogation results are dominated by the clear differences between arable crop and space crop (vineyards and orchards) farms concerning the needed pesticide amounts and the potential risks to the environment. The economic meaning of crop protection in relation to the total enterprise turnover is highest in space crop farms > big arable crop farms > medium arable crop farms, resulting in a corresponding interest in most aspects of the issue. The space culture farmers contributed most to the total number of mentioned ideas as well as to specific issues like application techniques, integrated farming, and organic farming. However, concerning the timing of applications, arable crop farmers were more active, indicating more flexibility in choosing optimal growth and weather conditions than permanent space culture farmers, who suffer from higher pest pressure. Permanent space culture farmers also have less flexibility in cropping areas: proposals in relation to buffer zones and protection of water bodies were mainly addressed by arable crop farmers. Space culture farmers regarded the political and economic influence on the possibilities of sustainable pest control clearly higher than arable crop farmers: the consumer conduct as marketing instrument is most important for direct food producers under high quality pressure.

2.5.2 Risk for ecosystems or environmental risks for consumers?

In 2005, a project was put to tender by the German Ministry of Education and Research BMBF, social-ecological research branch, to set up a risk realization, communication and management strategy for potential environmental risks that can hardly be assessed scientifically due to missing scientific knowledge. The proposal submitted by our consortium of agronomists, technology risk assessors, social-ecological consultants and ecotoxicologists from universities, small enterprises and research organizations decided to choose the issue of risks for chronic low dose effects on invertebrate communities by chemical use associated with animal production, e.g. of veterinary pharmaceuticals, food additives or biocides. A substantial part was the preparation of interactive education material for learning food web-interactions and the role of invertebrates in ecosystems. Whereas the poultry producers association indicated interest in participation, the project was not funded, because it was not regarded of public interest. Instead, projects dealing with environmental risks for consumers were funded.

2.5.3 Conclusion for the future

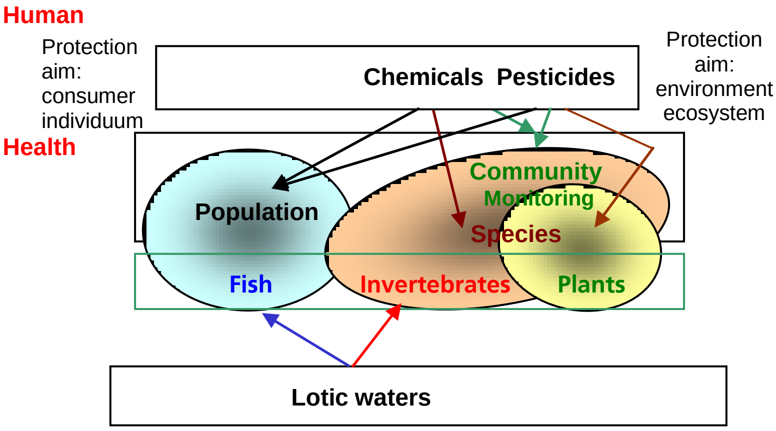
Our global economic system has to be considered when thinking about the maintenance or improvement of the environmental quality. The communication of benefits and risks of the use of chemical products has to be addressed to the consumers rather than to the producers and should be considerably improved. This should be in the responsibility of a panel of independent experts, i.e., organized by consumer protection agencies, which are highly accepted by the consumers. In contradiction to the current practice focusing on benefits and risks for consumers, ecological risks and benefits aspects clearly should be included. This might be supported by scientists of the different stakeholder groups, i.e. producers of chemicals, regulative authorities and NGOs. There is no sustainable development without communication of ecological needs to the consumer.

3 Water quality evaluation

3.1 Introduction

Ecotoxicological aspects of water quality evaluation can be related to substance evaluation when setting water quality objectives. This is mostly based on standard tests for aquatic toxicity.

Another aspect, needing more ecological background, is the identification of effects in the environment and attributing them to potential causes.



Water quality

The EC Water Framework Directive includes both aspects: By setting priorities for hazardous substances, the derivation of water quality standards is structured, delegated and accelerated. Chemical monitoring programs investigate whether the quality objectives for these substances are met or whether and how frequent critical values are exceeded as basis for the implementation of management measures. At the same time, water quality is evaluated independently from specific chemical water quality standards to guarantee the identification of hazards caused by other influence factors than by the measured concentrations of a very limited number of substances. There are two principally different approaches to detect hazards:

The indicator approach. Endpoints at different levels of biological organization serve as indicators for chemical stress. The indicators may be highly specific molecular biomarkers like VTG, or hormone receptor assays indicating sexual endocrine effects (see chapter 2.3.6), up to the assemblage of the site specific benthic community in streams and rivers, being used as indicator for the availability of oxygen and thus the quantity of organic pollution, associated with bacterial degradation and oxygen stress. When using indicators for water quality, the main challenge is to combine unspecific ones for a wide range of stressors to ensure that no potential hazard is missed, and highly specific ones for linking responses to effectors. However, the former tend to respond to a wide range of natural stressors, i.e., related to climate and reproduction cycles, the latter tend to only detect causes which are specifically looked for.

The habitat ecology approach. Aquatic communities are a central good of protection when evaluating water quality. Benthic communities are site-specific and integrate effects over time and causes. With improvement of wastewater treatment procedures, organic pollution as indicated by saprobe systems is no longer the only dominating problem of lotic water quality. Toxic chemical contaminants acting at much lower concentrations may now influence the development of aquatic communities and obstruct them in recovery. Modern approaches of water quality evaluation make use of habitat specific community structures by comparing them with the desired structure of comparable habitats of unpolluted streams. This desired structure has to be defined. Under influence of German culture, the definition has to be generic, as undisturbed water bodies do not exist in central Europe. The “Leitbild” is derived as potential ecological status of a water body. In Scandinavia, Spain and UK, the definition is more pragmatic and derived from reference sites of highest pristine status possible.

In the following, the main projects of the last twelve years dealing with ecotoxicological considerations of ecological water quality monitoring are shortly described, including conceptual projects for authorities, research projects funded by the agrochemical industry, and own projects supported by Fraunhofer:

- the comparison of different approaches to the classification of water quality in force at the time of the preparation of the water framework directive, with a specific focus on monitoring approaches (UBA project, chapter 3.2),
- the summary of the state of knowledge concerning monitoring studies as complied from a review and classification of pesticide related monitoring studies in Germany (BVL project) and the EPIF workshop (chapter 3.3),
- the conceptual planning, organization, supervision and evaluation of an extensive ecological monitoring study on the effect of pesticides in the field (IVA project, chapter 3.4),
- the representative investigation of distribution and fate of an adsorptive chemical in an outdoor artificial stream with considerations for potential exposure of different stream habitats (Fraunhofer project, chapter 3.5),
- and finally, as a kind of synopsis of the presented work, the concept of water quality evaluation as worked out for an EU 6th framework project (chapter 3.6).

3.2 Comparative assessment of national and international approaches to the classification of river health

In 1996, before the implementation of the water framework directive, water quality authorities and institutions of forty countries were asked for detailed information on executively applied classification concepts with focus on chemical water quality (ref. 29). The countries comprised 30 European countries (14 EU member states and 16 non-EU-member states) and ten countries outside Europe, regarded as important due to population size and high technical progress. 23 countries responded to this inquiry, the percentage of 56-60% was similar in all three groups. The transferred information was entered into 24 descriptions of different approaches and the consecutive comparison. Thorough assessment of the validity of the transferred and processed information (origin and character of the information, completeness, status of national implementation) resulted in 18 approaches with sufficient validity. The classification approaches were described in order to clarify and compare the aspects Reference/Leitbild; objective of protection; quality elements and parameters; status, level and derivation of water quality objectives; classification system and reporting of classified river health.

Understanding the different approaches within Europe and in comparison to abroad helps to understand the harmonization process, to accept national and EU-wide problems of implementing the water framework directive and identify potential for further development.

Leitbilder / reference sites. The expression “Leitbild” and its contents were only used in Germany and some neighboring countries. In all other approaches, the expression “reference” was used. In the Anglophone countries, Norway and Sweden, “reference” described nearly undisturbed sites being used as yardstick for the status assessment of sites under observation in monitoring programs.

Concerning all conditions out of the focus of the actual classification, reference conditions should be identical to that of the observed site. Concerning the conditions focussed on, reference sites should be pristine and represent the best achievable state of the classification scheme. This implements that for many local situations the reference can be different, depending on the resolution of the classification scheme.

A Leitbild is the theory-driven attempt to generalize conditions with the intention to derive natural conditions generically. Due to the lack of natural sites in the cultural landscape, this can either be attempted by relating to historical data (historical reference, as in The Netherlands) or based on analogy examples (potentially natural Leitbild, as in Germany, Suisse and Austria). As the Leitbild is an ideal set of conditions, it is outside the classification scheme. The best achievable status has to be derived and is called “operational Leitbild”. This can be restricted due to objectives of best water use.

In France, any orientation along “Leitbildern” or references was judged as handicap to the objective of classifications: The applied system was focused on water utilization levels (best use). This was also the case in most eastern European countries, in Finland and in Japan: The highest preferred goal of water managements was the achievement of the minimal water quality necessary for the most demanding utilization purpose.

Quality elements, objectives and requirements. When comparing the chemical water quality, only Canada and the Netherlands fully included the monitoring of dangerous organic and inorganic substances in water and sediment. In most of the other valid approaches, chemical monitoring was restricted to water samples, partly separating dissolved and particulate matter. In Scandinavia, monitoring was restricted to metal contaminations, however, in water, sediment and aquatic mosses. In UK, Ireland, Australia, Suisse, Austria, Belgium and Spain, the classification was mainly based on biological parameters, only supplemented by some general parameters and measurements of metal concentrations. Approaches mainly based on biological and inorganic parameters tend to be oriented at regionally or even locally different references, whereas approaches mainly based on dangerous organic substances tend to prefer a generalized or use-specific objective.

Water quality requirements, although mostly called standards, had very different legislative characters. They varied from obligatory standards over quality recommendations to reference values serving as orientation. Thus, the standard values differ due to their legislative status, the obligatory standards mostly being the highest ones.

Water framework directive. The overview of the actual state of the national approaches, the possibilities and shortages, enabled an assessment of the compatibility to the European concept.

The orientation by reference sites was already in use in UK and Scandinavia. The main good to be protected is “aquatic communities” and the objective of a good ecological state was inherent in the approaches of UK, Spain, the Benelux-States, Scandinavia, Germany and Austria.

The biological parameter values have to be compared with appropriate values of reference sites. For this, approaches from UK, The Netherlands and Austria contained the most promising solutions. The indicator approaches being used in most European countries are now less important.

The derivation of chemical water quality objectives is similar to the German approach. The concept of classification and presentation reflects the average of the European approaches.

As a whole, the conception of the water framework directive took up actual developments and transformed them as consequently and directly as possible into a harmonized framework.

3.3 Monitoring effects of pesticides in the field: overview of concepts, objectives and applications

Review of German monitoring projects. When focusing on monitoring of pesticides in the field, the objectives are moving towards the refinement of a substance related risk assessment. In this context, monitoring studies and field studies with an experimentally integrated exposure sometimes are merged. In a review of the most important published monitoring studies in Germany aiming at the effects of pesticides on non-target organisms ref. 55) we tried to classify objectives, approaches and methodologies and evaluate the reported effects with respect to the influence on the pesticide notification process. Not included were field studies for pesticide notification and routine monitoring of rivers. The main work was performed by Udo Hommen.

In all, 41 studies were analyzed and classified according to

- the investigated compartment
  - water body
  - terrestrial structures (mostly off-crop)
  - soil (mostly in-crop)
- the methodology
  - chemical monitoring of active substances (total or bioavailable concentrations, the latter also depending on media properties, e.g. pH, DOC, particle content)
  - ecotoxicological monitoring, e.g. active biological monitoring (introducing encaged or sessile organisms) or conducting bioassays with samples taken from the sites; investigations of molecular or physiological biomarkers
  - ecological monitoring, e.g. observation of effects on populations (population densities and structures) or species (spectra and abundances)
- the main objective
  - water or soil quality (4 studies)
  - influence of agriculture (3 studies, mainly without analysis of pesticides)
  - crop specific risk assessment (7 studies)
  - active substance related risk assessment (4 monitoring and 9 field studies)
  - exposure (mitigation) of non-target areas (6 monitoring and 7 field studies)

Most of the analyzed projects investigated pesticide fate and/or effects in water bodies, chemical monitoring being the main methodology. There were no findings of studies on pesticide field effects on vertebrates.

Ecological monitoring was performed with highly different intensity regarding the number of sites and frequency of sampling. It was partly added by ecotoxicological monitoring (bioassays).

Approximately one third of the studies investigating effects of exposure following good agricultural practice (n = 34) showed no effect at all. One third revealed only slight or temporal effects with observed recovery in due time, which is regarded as acceptable for higher tier studies within the pesticide registration procedure. The last third exhibited effects with unclear or without recovery within one year. In aquatic systems, the effects on the population dynamics were mostly temporal ones. However, lasting effects on the structure of aquatic communities were regarded probable in the orchard culture of the Altes Land at very low distances to the crop ((ref. 17), see chapter 3.4) and in streams of the Braunschweig region with high potential of run-off exposure (Liess and Schulz 1999, Liess et al. 2001). Terrestrial non-target arthropods in off-crop areas exhibited only slight, if any, direct effects. However, there are some hints of long-term effects of insecticides to locusts (Kühne et al. 2001, Tristsch et al. 2001). Up to now, there are no German studies showing a clear effect of herbicides on the vegetation of non-target areas. The observed shift to grass dominated communities (Jüttersonke et al. 2004, Roß-Nickoll et al. 2004) could not be separated from other influence factors such as nutrient intake. Two BBA studies identified long-term effects on in-crop soil algae and Collembola.

Whereas in-crop effects are a matter of cost-benefit-analyses, can be tolerated to a certain extend and should be minimized by appropriate selectivity of the pesticide, off-crop effects in non-target areas like water bodies should be avoided by risk mitigation measures such as untreated buffer strips or drift-reducing application techniques.

The environmental risk assessment of pesticide use is based on testing the active substance and the formulation in standard single species tests and, if necessary, in higher tier tests and/or field studies. All these studies focus on effects caused by the single product, excluding all other stressors to enable a powerful evaluation of the test item. Under field conditions, however, populations and communities can be affected directly or indirectly by a combination of different pesticides as well as by other agricultural activities such as fertilization, plowing or mowing, and by habitat characteristics as diverse as vegetation, particle size or connectivity. These influence parameters are of different importance for each species, cause high variability and mask potential effects of pesticides. Thus, a detection of clear pesticide effects is difficult and only feasible by relating them to real exposure. This includes the observation of peak concentrations and the time profile of exposure and represents a challenge when applying it to a sufficient number of sites that is necessary for a statistically powerful effect evaluation.

Monitoring studies can be useful for checking the realism of concentration and effect assessments, for identifying differences in scenarios, which may result in scenario-specific risk mitigation measures, and for the identification of sensitive indicator species, which may be used in higher tier studies to represent sensitive communities in regional risk assessments.

Chemical monitoring is useful for substance-specific objectives such as for surveillance of risk mitigation measures and the detection of misuse. The main problems are the sampling for detection of peak concentrations, the coverage of diverse groups of substances with different work-up and detection methods, and the warranty of sufficiently low limits of quantification in relation to the expected NOEC. Consequently, the results of a chemical monitoring can often only be regarded as minimum exposure. For the derivation of risk potentials from chemical monitoring results, the findings for individual substances can be compared with toxic threshold concentrations from pesticide notification studies. A different concept is the use of toxic units by relating the measured concentrations of the individual compounds to a specific toxicity endpoint, e.g. the EC50 for *D. magna,* and summing up these quotients. Both approaches need the access to recently actualized ecotoxicity data. A publication of the most important ecotoxicological endpoint data used in the (national) registration would facilitate the use of chemical monitoring data for risk assessment rather than only for exposure assessment and enhance the comparability of data evaluations. For applying the toxic unit concept, the EC50s for Algae and Daphnia could be used to normalize exposure to different substances on effective concentrations for macrophytes and invertebrates.

The problem of ecotoxicological or active biological monitoring is the unclear relevance of the indicator organism fort the local community. In case of high sensitivity, this kind of monitoring can be used as an early warning method or a trigger for further investigations. As addition to ecological or chemical monitoring, the methodology has its value as interpretation tool for the separation of pesticide effects from those of other influence factors.

Regarding the observed endpoints on the population and community level, ecological monitoring generates the most relevant data for the risk assessment of pesticide use in the field. The main problems are the separation of pesticide effects from those of other influence factors, the potential development of communities adapted to pesticides and the finding of appropriate reference sites representing the untreated control situation. As a recommendation, a monitoring study should contain a sufficient number of sites with very similar conditions but different pesticide exposure for a potential evaluation by regression analysis. Bioassays or biomarkers can be used to separate responsible influence factors. References sites for the agricultural landscape have to be found and generally accepted to represent species spectra of relevant and realistic communities to be protected.

As a general conclusion from this review, thorough ecological monitoring is recommended in regions with high exposure potential, such as areas with a high intensity of pesticide applications (e.g. orchard cultures), with high run-off potential, or small catchment areas with intensive agriculture. In areas with special regulation, i.e., the „Altes Land“, the obligatory chemical monitoring should be supplemented by an ecological monitoring of sufficient sites to enable the separation of pesticide effects from that of other influence factors. When interpreting monitoring data it is helpful to include results of notification studies, especially higher-tier studies. A potential source could be the Alterra expert system PERPEST, which makes use of micro- and mesocosm studies to predict classified effects (Brock et al. 2000) on aquatic communities (van den Brink et al. 2002). However, data from the notification procedure are related to one product. The toxic unit concept can only be regarded as a very simple model to assess mixture effects. For assessing effects of culture-related pesticide use patterns, community level studies on the effects of those patterns are a clear improvement. First results were reported for typical application patterns in potatoes and tulips (van Wijngaarden et al. 2004).

Due to the individual character of each monitoring study, a repetition is not possible. To achieve a performance according to the best practice, a guidance document on the performance of monitoring pesticide effects in the field including quality criteria would be helpful. The conclusions of the EPIF workshop could be a starting point.

EPIF workshop. During the EPIF workshop held in Lacanau, France, the importance of defining accurately the exposure was identified as an important issue for regulatory purposes (Liess et al. 2005). Indeed if a link between chemical causes and biological effects can be established in the monitoring studies, then it may be possible to examine the correlation between critical effect concentrations detected in the field and those predicted from estimated exposure under Directive 91/414. However, the fact that a correct quantification of the time profile of exposure is even difficult in controlled field studies and nearly impossible in monitoring projects, was not sufficiently acknowledged.

Another important outcome of the workshop was that results from monitoring studies highlighted issues that would be relevant for increasing realism in future risk assessment. For example, the following parameters or issues should be considered and/or implemented, as they are or could be important to assess the occurrence of effects under field conditions.

- Chemical monitoring: This can be effectively used, for example, in surface waters for regulatory purposes to show that risk mitigation measures work properly. However, to date the prediction of effects based on such data is only approximate, because many processes can influence biological effect and their recovery. *Note: Uncertainty is added by the sampling scheme for detecting exposure peaks and profiles.*
- Route of exposure: The relative importance of different routes of exposure should be appropriately implemented in regulatory risk assessment. For example, studies on birds pointed out that the role of dermal exposure is likely to be underestimated in current regulatory risk assessment schemes.
- Mixture toxicity: Some results showed that risk prediction based on the most toxic product should lead to reasonable regulatory decisions. However, further research is needed to determine whether risk assessment that considers only single compounds is generally conservative for mixture toxicity.
- Target images: most of the issues cited above should help to define target images that are needed to evaluate monitoring studies. Close contacts between monitoring activities

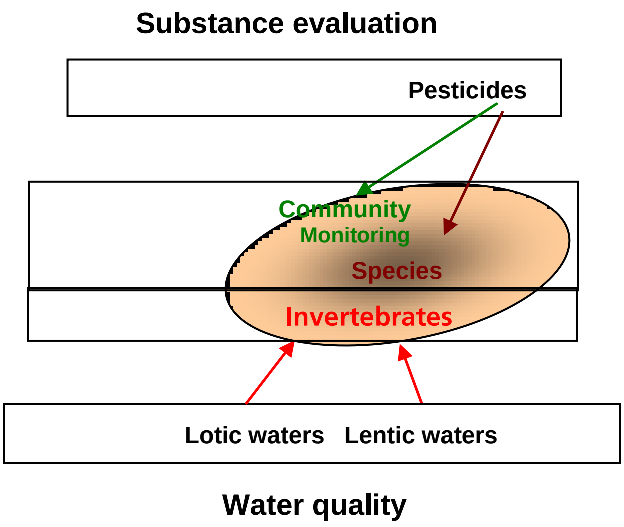
under Directive 91/414/EEC and the Water Framework Directive are needed so that overlaps can be avoided.

- Environmental parameters: They may increase or decrease the sensitivity of organisms towards pesticides. Such parameters should therefore be included in risk assessment.
- Biological and ecological data: more data are needed on the different biological and ecological traits of non-target species. For example, these data could be included in the information on a specific landscape and analyzed using GIS systems to facilitate use for regulatory purposes and to improve the realism in risk assessment.
- Indirect effects: They have been observed for all groups of organisms and their relevance as regards the long-term repercussions on populations and communities should be investigated. The causal relationship to direct effects could be better investigated by, for example, the use of biomarkers.
- Recovery and recolonization: these processes have important implications for assessing the occurrence of effects. However, it will be challenging to include these concepts into a risk assessment in a field-relevant manner. Indeed, repeated exposure or other stressors might disturb these processes. Furthermore, species with a low recovery potential (i.e., univoltine species and/or species with low dispersal potential) should be considered more carefully than species with high recovery potential (i.e., short generation time and/or high dispersal ability). *Note: The occurrence of species with low recovery potential can be regarded as integrative evidence that no relevant effect has happened in the recent past.*
- Incidence schemes: The organization of such schemes should be established or improved when monitoring studies are difficult to conduct (i.e., for vertebrates such as birds). The schemes are a safety net for the regulatory system and can highlight the need for appropriate risk management and issues regarding misuse and abuse.
- Landscape analysis: Landscape characteristics can modify the importance of field effects of pesticides and recovery (e.g., conservation headlands support a diverse community of terrestrial organisms in agricultural landscapes; forested stream sections facilitate the diversification of aquatic communities through recolonization). Classifying landscape characteristics and including habitat quality in risk assessment may put the risks of contamination into context with respect to other stressors.

It was generally accepted within the workshop that it is challenging to determine pesticide exposure as the cause of effects in the field. This difficulty is due to problems of natural variability, multiple substances and multiple stressors, confounding factors, and insufficient statistical power. Concerning the current regulatory risk assessment procedure, many participants felt that current approaches are reasonable to ensure the protection of non-target organisms and, in some cases, may be considered too conservative. However, effects of pesticides were identified in several of the field studies presented. It could not be determined whether good agricultural practice or misuse was responsible for the observed effects in many of these studies; therefore, some uncertainty remains about the actual level of protection. It was stated that further research is needed to evaluate accurately the degree of protection, so that risk assessment procedures can be adjusted if necessary.

3.4 Monitoring effects of pesticides in the field: Results of a research project funded by the IVA

3.4.1 General scope and approach



Ecological monitoring programs for aquatic systems are often designed to measure the community structure using macroinvertebrate and/or fish species. However, they have rarely been used so far to analyze the significance of pesticide-specific effects in relation to the overall impact of the agricultural land use. Aquatic ecosystems in agricultural landscapes are frequently exposed to different kinds of **stressors that may cause changes in** biological community structure. As any kind of land management for agricultural purposes has ecological implications, it is difficult to quantify the significance of individual stressors in relation to overall changes of the agricultural environment. Strong causal relations can be made for stressors with well-established direct and/or indirect effects, such as nutrients and pesticides. However, landscape-related variables, such as the type of terrestrial vegetation, the topography, geomorphology and hydrology, are important as well, because of their known influence on the aquatic habitats and species mobility.

Benthic macroinvertebrate communities are widely used as indicators for water quality. In contrast to chemical water monitoring programs or the monitoring of organisms with short generation times (e.g. plankton), they integrate potential impacts over considerable time periods. Compared to fish, macroinvertebrate sampling is easier and, because macroinvertebrates are less mobile than fish, effects can be better related to local exposure. In the UK and Australia, RIVPACS (River Invertebrate Prediction and Classification System, Wright et al., 2000) is used for classifying rivers according to their habitat characteristics for the benthic macroinvertebrate community, and for predicting the community of a monitoring site. The system is based on environmental and physical characteristics of a specific site, such as longitude, latitude, altitude, distance from source, water width and depth, substrate composition, discharge category (or median current velocity), alkalinity, slope and air temperature. For the purpose of the presented study sponsored by the german agrochemical association IVA, we amended this list of well-established influence variables on macroinvertebrate community structure with variables specific for pesticide exposure. Several factors are expected to have a large impact on the macroinvertebrate community structure: habitat characteristics, climatic conditions, and anthropogenic operations such as water bank management, weirs etc., and especially in agricultural areas fertilizers, soil cultivation (erosion) and pesticide exposure. While some parameters act more constantly, others are more temporally limited events. In the latter cases resilience (tolerance, sensitivity), recovery (from within the impacted region) and recolonization potential (from outside the impacted region) are important factors influencing the community structure as shown in Figure 39. We used this figure to analyze the causality of potential pesticide exposure for changes of the macroinvertebrate community structure in two study areas:

- The “Altes Land” near Hamburg is characterized by intensive orchard cultures within a close web of ditches, representing a worst case of exposure to aquatic communities. The ditches were originally built to drain the land for the orchards. They are of relatively equal morphology and highly connected, reducing the overall variability of the macroinvertebrate community, a general requirement for linking cause and effect and separating stressor influences from natural variability. From a methodological point of view, the publication of this part of the study (ref. 17) is a key publication for the setup, performance and multivariate evaluation of targeted monitoring studies.

- The Braunschweig region is characterized by intensively cultured arable land in a diverse landscape, both aspects being more representative of situations in central Europe.

For each agricultural region, 40 sampling sites were selected at different ditches to fulfill the following general requirements:

- Lowest possible variability in habitat characteristics, such as stream morphology, substrate, aquatic macrophytes, plant cover on ditch banks
- Widest possible range of potential pesticide exposure attributes, such as land use, crop protection strategy and intensity, distance of crop to the water body.

The sites were sampled five times for macroinvertebrates, water quality and other habitat characteristics (Table 14) in autumn 1998, spring 1999, summer 1999, autumn 1999, and spring 2000. Macroinvertebrates were collected from sediment substrates (Braunschweig) or macrophyte and sediment substrates (Altes Land, pooled for evaluation). For most individuals, taxonomic determination was possible down to the level of species or at least genus (ref. 41), 42)).

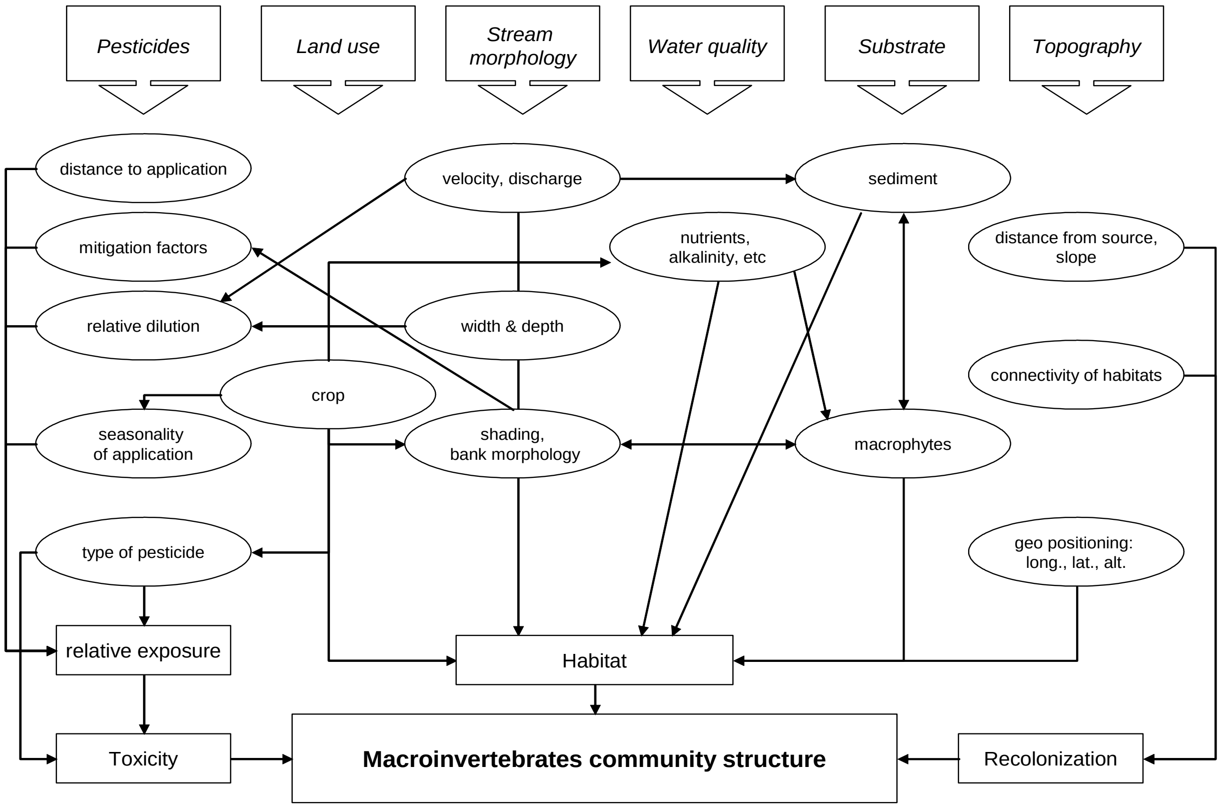


Figure 39: Factors driving macroinvertebrate community structure in agricultural landscapes
Habitat characteristics, toxicity, recovery (from within the impacted region) and recolonization potential (from outside the impacted region) are important factors influencing the community structure (see boxes at the bottom). These main factors depend on several influence variables, which can be grouped in ‘pesticides’, ‘land use’, ‘stream morphology’, ‘quality of water and substrate’, and ‘topography’ of the landscape (arrows at the top). Each of these groups can be described by different variables, which may be related to each other in a conceptual way (ellipses and arrows).

Table 14: Environmental factors determined for all sites
(AL) "Altes Land"; (Bs) Braunschweig region

Water quality (each sampling)	**Structure & morphometry (each sampling)**	**Structure & morphometry (once)**	**Exposure related factors (twice, based on autumn 1998 and summer 1999)**
• Temperature • Dissolved oxygen • pH • Conductivity • Alkalinity • total N; NH_4; NO_3; NO_2 • total P; PO_4 • Mg; Ca • Turbidity	• Water surface width • Water depth • Macrophyte coverage • Shading • Bank vegetation • Detritus • Anoxic sediment	• Width of ditch • Slope of ditch bank • Distance to main ditch • Sediment structure	• Potential for exposure • (AL) only drift depending on distance dependent drift rate, width of the ditch and water depth (Bs) drift ad runoff including slope and soil surface, upstream cultures and water discharge • Distance to treated cultures

The similarity of some variables (Table 14) was very high between the sites, especially in the "Altes Land", due to the low variability in the ditch morphology. Such non-differentiating variables are of limited value for the further evaluation and were neglected to improve the analysis. The decision logic used to evaluate the data followed a systematic approach reducing the variability of the dataset and strengthening causal relations with each step:

1) *Identification of outliers*

Sites were identified as outliers, if conditions changed considerably during the study (e.g. change of land use, change of morphology, dry-off, presence of H_2S). A Two-Way Indicator Species Analysis (TWINSPAN) was used for a first classification of the sites in order to prove whether the outliers identified by expert judgment indicated different communities compared to the other sites. TWINSPAN is widely applied by ecologists for classifying species and samples, producing an ordered two-way table of their occurrence. It uses a hierarchical approach by successively dividing samples and species into categories. The analysis was performed by using the sum of the abundance data over all sampling dates for each site.

2) *Influence of land use or geography on community structure*

A TWINSPAN of the community data without the outliers served to test whether sites with different land use or geography would fall into different clusters. This classification was amended by ordination techniques to analyze the importance of the different environmental factors. Depending on the distribution of data, Correspondence Analysis (CA) or Principal Component Analysis (PCA) and Redundancy Analysis (RDA) were applied. The environmental factors were ranked according to their correlation with the first two axes of the CA or PCA and to the part of total variance that they were able to explain in a CCA or RDA, using forward selection of factors.

3) *Influence of exposure potential on community structure*

In step 2, sites were grouped to dominant influence factor regimes independent from pesticide exposure. Thus, the analyses in step 3 were performed individually for each cluster. Again, CA or PCA and RDA or CCA were applied to identify the important environmental factors for each cluster of sites. As the evaluation is aimed at the relative importance of pesticide exposure, only clusters containing sites with considerably different potentials for exposure could be analyzed.

The number of taxa as a simple community index was plotted over the log transformed potential for exposure, and the relationship was tested by linear regression. In addition, the species composition was analyzed qualitatively by identifying species missing at sites with a high potential for pesticide exposure and comparing the presence of species listed by the relevant regional nature conservation board list, the "Red List" of Lower Saxony, in the different site groups.

4) *Importance of exposure potential for seasonal dynamics of community structure*

If in step 3 the exposure potential was identified as one of the most important factors driving community structure (only in the "Altes Land"), the data were analyzed for the influence of the exposure potential over time. For that purpose, Principal Response Curves (PRCs, Van den Brink and Ter Braak 1999) were used. The sites along orchards were grouped into sites with a low, medium and high potential for exposure (4, 10 and 7 sites, respectively). PRCs were also calculated without the sites along organically farmed orchards.

3.4.2 Project part "Altes Land"

In the "Altes Land" (ref. 17), 48)), pesticide applications during the investigation period followed the recommendations made by the local agricultural support authority (Obstbauversuchsring des Alten Landes e. V. in Jork). To be able to detect clear pesticide impact, a deviation from good agricultural practices was tolerated releasing the farmers from mitigating label restrictions during the study period and allowing them to treat all rows of trees, irrespective from the distance to the water bodies. The choice of ditches and the sampling and determination of organisms was performed by our cooperation partner planula (Michael Dembinski and his team), Hamburg-Altona.

According to the type of agriculture on the neighboring fields, the following groups of ditches were distinguished: grassland (G, 13 sites), unused old apple orchards (U, 3 sites), organic apple culture (O, 6 sites), and integrated apple culture (I, 13 sites). At five sampling sites, land use was different along both banks (one site U-I, one site U-G, three sites O-I). The term "integrated" indicates orchard management including the use of synthetic organic pesticides, while "organic culture" refers to plant protection without synthetic organic pesticides. However, copper-based fungicides, other inorganic substances, and natural pesticides like pyrethrum and biological agents (*e. g. B. thuringiensis*) were allowed.

1) Identification of outliers

During the sampling campaigns, nine sites were identified as outliers: Five sites became overgrown or dried up during the study period (sites U-22, OI-26, O-27, O-28, G-32). At two sites the farming of the neighbored area changed or the ditch was drained (sites G-37, G-39). One ditch was much larger compared to all other ditches (site O-30), and one ditch showed an extreme variability of water levels (site I-35). We used TWINSPAN to conduct a classification of the 40 sites based on the species data only (Figure 40). Most of the sites identified as outliers were clustered in one group, indicating that these sites showed macroinvertebrate communities differing from the remaining sites. The outliers were not considered further. In the samples of the remaining 31 sites, in total 23703 individuals of 181 taxa were found.

2) Influence of land use on community structure

A TWINSPAN analysis with the remaining 31 sites produced a first subdivision between eight of the ten grassland sites and all the other sites (Figure 40). The mollusks *Pisidium milium*, *Hipeutis complanatus*, *Sphaerium corneum*, *Bathyomphalus contortus*, and midges from the group Chironomini were identified by TWINSPAN as 'indicator taxa' for the grassland sites.

The importance of the neighborhood (grassland or orchard) was differentiated from other environmental factors by ordination techniques including the potential for exposure. A PCA confirmed the TWINSPAN result indicating that communities at sites along grassland differ from those at sites along orchards. When correlating the environmental factors to the first two ordination axes, the highest correlation with the first axis was given for the distance to trees and the factor land use (grassland "yes" or "no"), respectively, while the exposure potential and macrophyte coverage correlated best with the second axis. Thus, community patterns can be explained best on a first level by land use, while the exposure potential seems to be the most important environmental factor for sites along orchards. In an RDA using automatically forward selection of factors based on the proportion of total variance explained by the respective factor, also the distance to the trees was identified as the most important factor. To avoid mixing effects of trees close to a ditch from effects of the (correlated) potential for pesticide exposure, it was decided that communities in ditches within grassland should not be used in the following analyses.

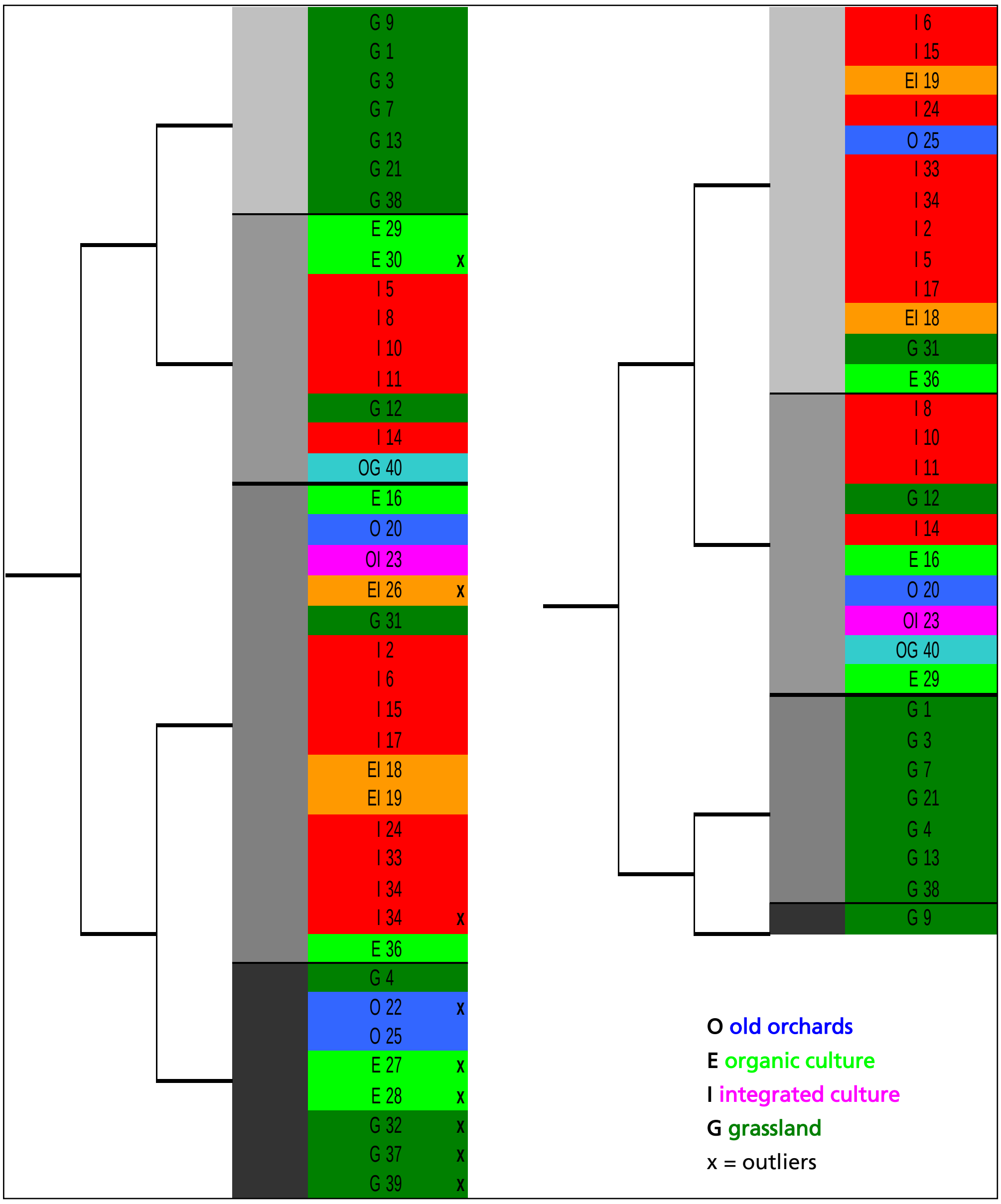


Figure 40: TWINSPAN grouping of sites the “Altes Land”, based on added abundances over all sampling dates
Left: Analysis with all sides, right: analysis excluding the outliers. The first separation divides most of the grassland sides from all others including untreated orchards.

3) Influence of exposure potential on community structure

After the omission of land use (grassland or trees adjacent to the ditches), the estimated potential for exposure was identified as the most important environmental variable explaining the community structure variability for the 21 remaining sites along orchards. The exposure potential was the only factor with a highly significant correlation ($p<0.001$) to the first axis of a PCA, while correlations to distance to trees, calcium concentration, total phosphorus concentration, slope of the bank, pH, and depth of the water body were significant on the 5%-level. Width of the water body, oxygen and phosphate concentrations were the factors, which correlated significantly with the second ordination axis of the PCA. Also in the canonical ordination (RDA), the exposure potential was the factor explaining most of the variance.

The total number of taxa was generally lower at sites with a high potential of exposure and negatively correlated with the log-transformed exposure potential ($p < 0.001$, Figure 41). For example, Bivalvia were totally missing in the high exposure potential sites, and endangered or vulnerable species according to the Red List (Figure 42) were only found at sites along grassland or in the low or medium exposure group (for grouping see Figure 43). In the latter group, these species were found in ditches along organic as well as integrated apple cultures.

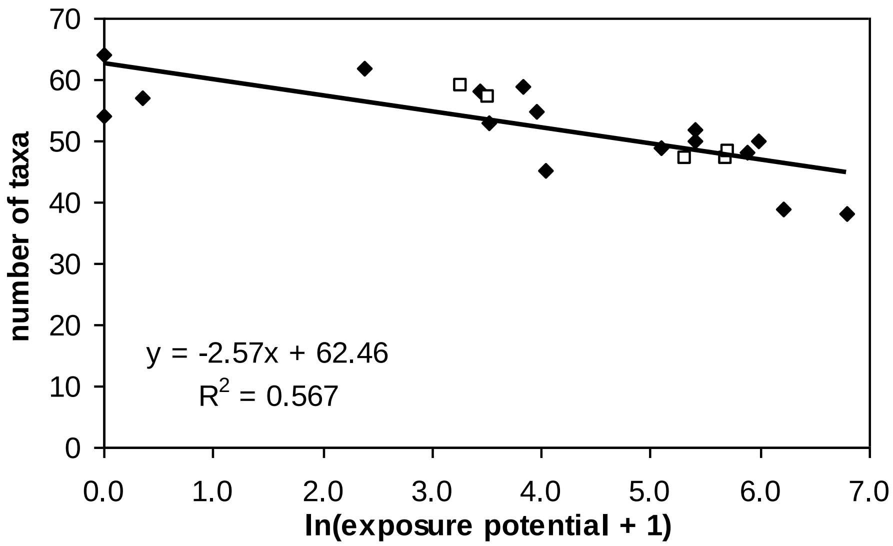


Figure 41: Regression of the number of species found in the five samples per site on the log-transformed exposure potential
(sites with organic orchards at one or both sides are indicated by open squares).

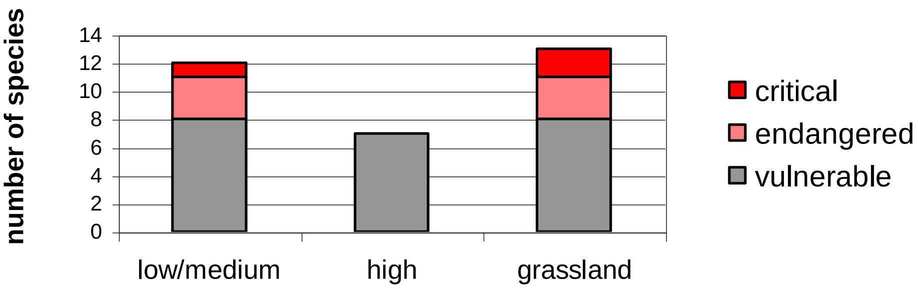


Figure 42: Presence of species listed in the Red Data book of lower Saxony
For graphical presentation, sites with low and medium potential for exposure were combined to get a comparable number of sites in each group.

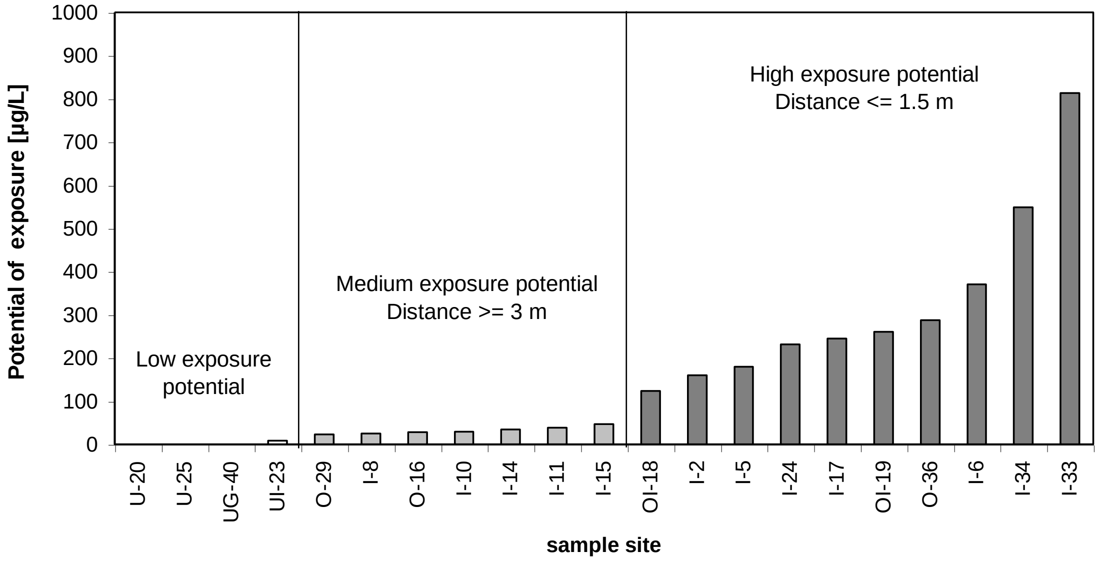


Figure 43: Range of estimated potential of exposure and grouping for PRC analysis
No exposed sites were available in a distance between 1.5 m and 3 m from the first row of trees, since the first row was planted either close to the ditches or far enough for a tractor lane.

4) Importance of exposure potential for seasonal dynamics of community structure

The results presented so far were obtained for the species data summarized over the five samplings from autumn 1998 until spring 2000. In order to analyze seasonal differences in the relation of community structure to exposure potential, PRCs were calculated for the "medium" and "high" exposure groups using the "low" group as a control (Figure 43). Because no clear differences between organic and integrated sites had been indicated so far, the first PRCs were calculated for the 21 sites along orchards including organically farmed ones (Figure 44). When comparing the sites of the "medium" group with those of the "control" group, only small deviations occurred which were never found to be significant (Monte-Carlo permutation tests). However, for the sites with a high exposure potential, the PRCs indicated larger deviations, which were significant for four of the five sampling dates. The species weights in the right part of the graph provide information about the relative importance of the different species within the PRCs. Species with a high positive weight show a similar pattern as described by the PRCs, contributing most to them, while species with a high negative weight exhibit the opposite dynamics.

To focus the comparison on sites with a comparable treatment pattern, additional PRCs were calculated without the five sites neighbored by organically farmed orchards at one or both sides (Figure 45). Differences of the "medium" to the "low" group are more pronounced now; they are significant in autumn 1998 and in summer 1999. Differences in the "high" exposure group did not change considerably. The species weights were ranked in a similar way as in the PRCs including the organic sites, indicating strongest effects for *Asellus* and *Cloeon* and an opposite pattern for most of the Pulmonata species.

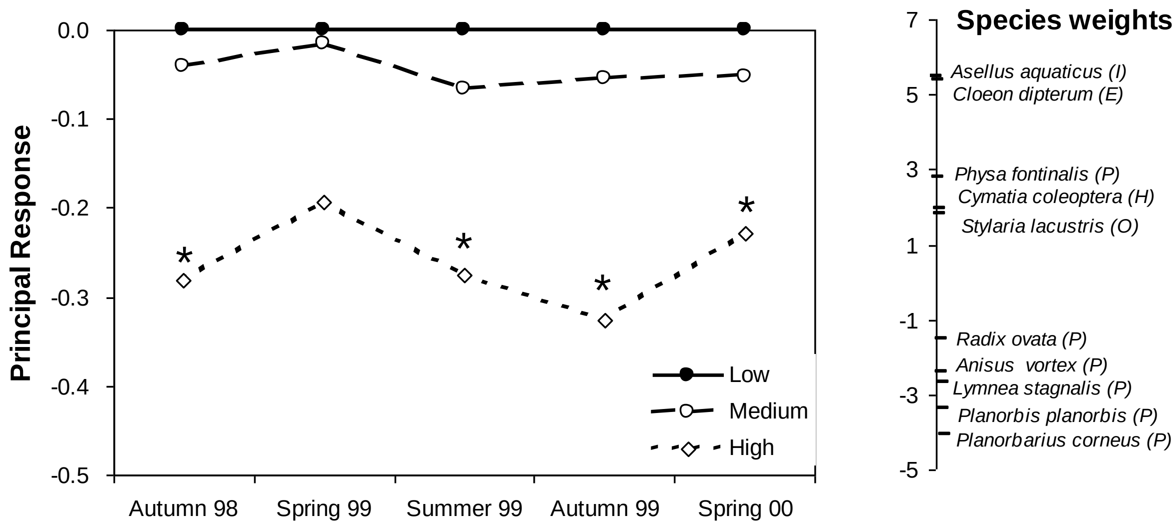


Figure 44: Principal Response Curves for the 21 sites along orchards, including organically farmed ones
Quality of the PRC: 10% of total variance explained by time, 15% of total variance explained by treatment, 42% of the variance explained by treatment captured by the PRC, p-value for PRC in total = 0.006; * indicate significant ($p<0.05$) differences to the low exposure group according to a redundancy analysis including permutation test and using the log-transformed potential for exposure as environmental factor; species weights only shown for the ten species with the highest and lowest weights (I = Isopoda, E =Ephemeroptera, P = Pulmonata, H = Heteroptera, O = Oligochaeta)

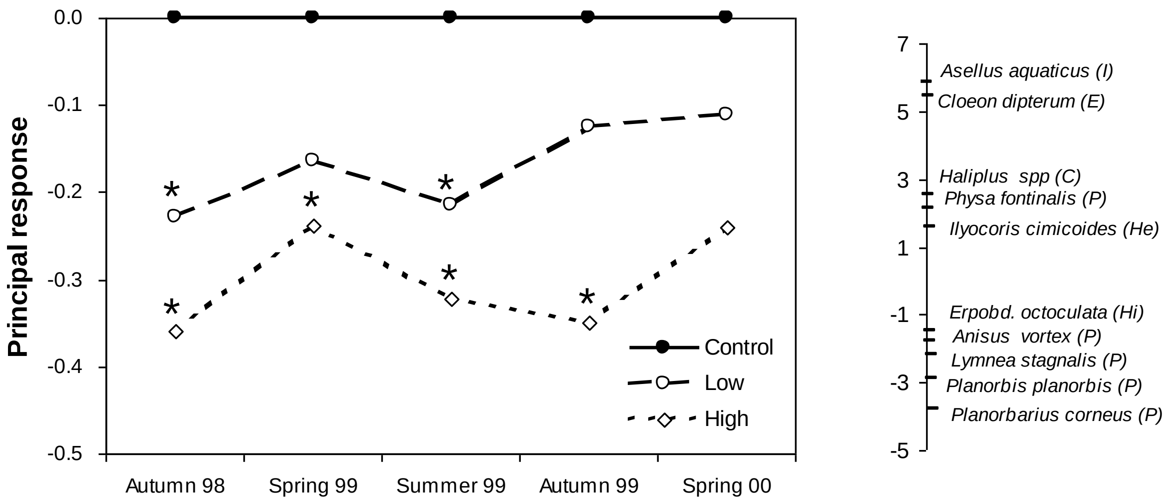


Figure 45: Principal Response Curves for the sites along orchards (without organically farmed ones, n=16)
Quality of the PRC: 12% of total variance explained by time, 18% of total variance explained by treatment, 37% of the variance explained by treatment captured by the PRC, p-value for PRC in total = 0.036. * indicate significant ($p<0.05$) differences to the low exposure group according to a redundancy analysis including permutation test and using the log-transformed potential of exposure as environmental factor. Species weights are only shown for the ten species with the highest and lowest weights (I = Isopoda, E =Ephemeroptera, P = Pulmonata, H = Heteroptera, O = Oligochaeta).

3.4.3 Project part "Braunschweig region"

The presented methodology was also applied to the generation and evaluation of monitoring data in the Braunschweig region. Due to a much more variable landscape and use situation, the selection of sites was challenging. The choice of sites and sampling and determination of organisms was performed by our PhD student Stefan Pantel in cooperation with Matthias Liess, Braunschweig. As we were not able to find 40 sites meeting the general requirements in a close area (it was intended to find them in the loess region in the south of Braunschweig), we had to extend the search. The finally sampled sites consisted of

- 20 sites south of Braunschweig (loess region)
- 12 sites north of Braunschweig (heath region)
- 8 sites west of Braunschweig (Weser mountain region)

1) Identification of outliers

During the sampling campaigns, 15 sites were identified as outliers. Most of them dried up during the study period, especially in the very dry summer and autumn of 1999. This effect mainly occurred in the sandy heath region (Braunschweig north), where nine of twelve sites were affected. The data of the first sampling campaigns were used for an identification of the influence of geographical regions on the community structure.

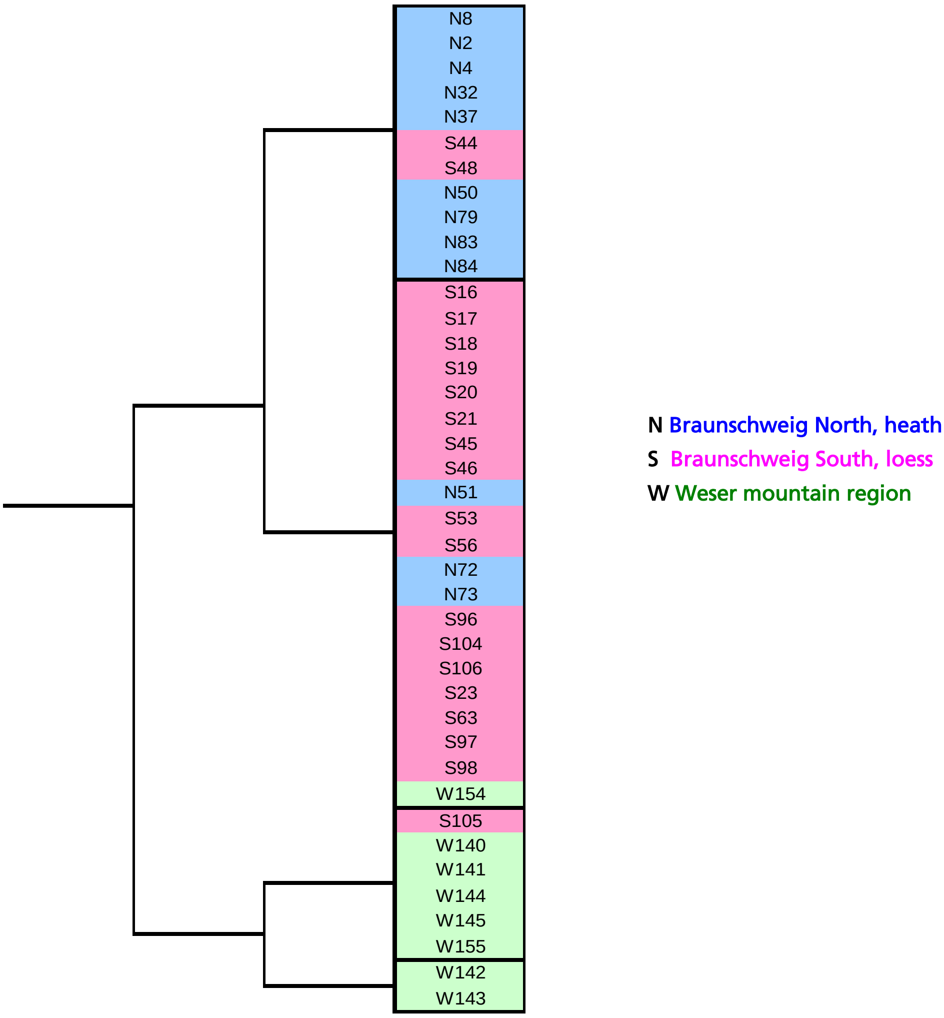


Figure 46: TWINSPAN grouping of sites the Braunschweig region, based on added abundances over the first sampling dates
The first separation divides the mountain sites from all others, the second level divides most of the heath from most of the loess sites.

2) Influence geographical subregions on community structure

A TWINSPAN analysis with all sites produced a first subdivision between nearly all Weser mountain sites and all the other sites (Figure 46). The second separation level produced one cluster comprising 75% of the heath and only 10% of the loess sites, and another with 85% of the loess and 25% of the heath sites.

The importance of factors related to the specific geographical subregions was differentiated from other environmental factors by ordination techniques including the potential for exposure. A CCA confirmed the TWINSPAN result indicating that the sites belonging to different geographical subregions represent different communities (Figure 47). The relation to the environmental factors consistently explains the preferences for species forming the communities in the subregions. While the R-value, indicating the geographical longitude, just hints to the western position of the most different communities of the Weser Mountains, the H-value, indicating the geographical latitude, indicates to the special communities of the northern heath sites. Highly differentiating properties are alkalinity, mainly spanned by the loess sites containing most nutrients, and the stream velocity, mainly spanned by the mountain sites. The sediment structure is also characteristic. While stones and gravel are correlated with community differences in the mountain sites, sand is dominating the heath sites. Mud, loam and loess are directing to the loess sites, but show weaker correlation to community structure. Besides these geographical factors, depth and width of streams and ditches are very important for the community structure. As community patterns can be explained best by geographical subregions, it was decided that communities of the different subregions were investigated separately to avoid mixing regional effects from effects of the (correlated) potential for pesticide exposure.

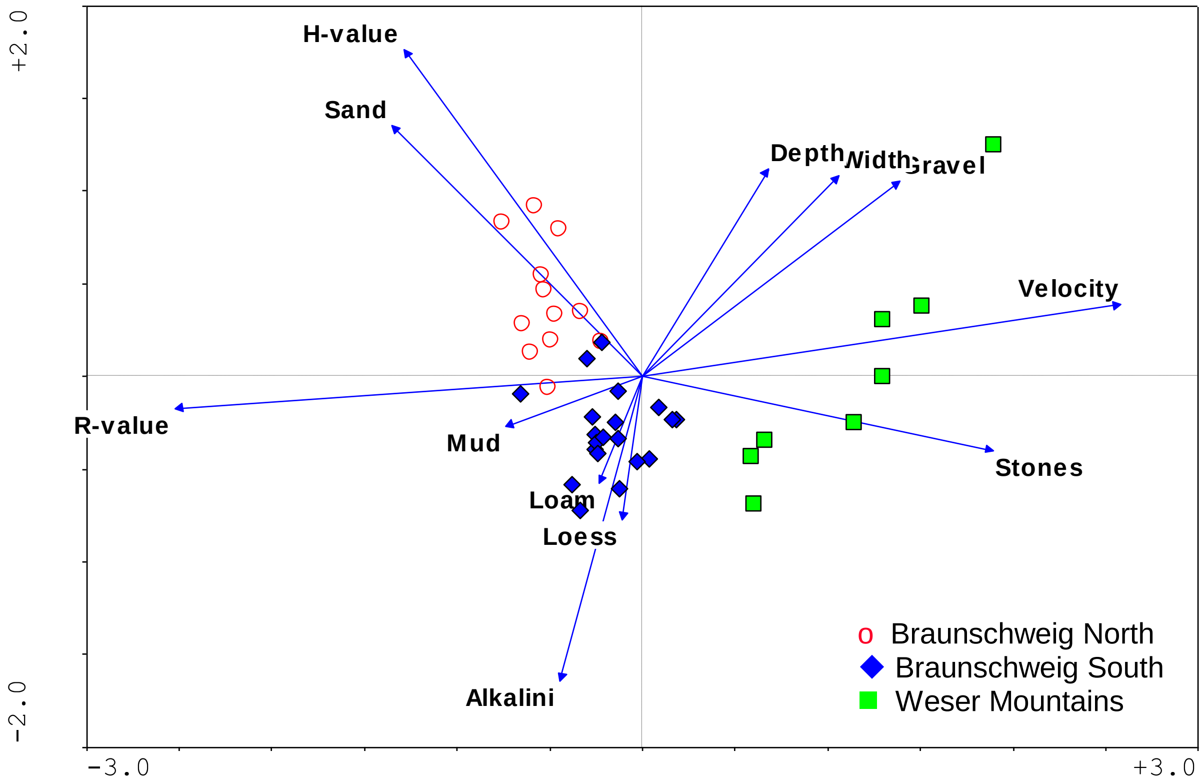


Figure 47: A Canonical Correspondence Analysis (CCA) grouping sites the Braunschweig region according to different environmental factors, based on added abundances over the first sampling dates
The length of arrows indicates the goodness of correlation. The analysis explains 15.4% of species data variability, p(first axis) = 0.005.

3) Influence of exposure potential on community structure

Grouping of sites according their potential for pesticide exposure revealed that Weser mountain sites were neither exposed to spraydrift nor to runoff in a sufficiently different way and had to be excluded from further evaluation (Figure 48). Braunschweig North and South sites in principle were well selected with respect to the range of potential exposure. However, the high losses of heath sites by drying up (75%) reduced the analysis to the first year and enabled a proper analysis of the full data for the loess sites only. When analyzing effects of the potential for pesticide exposure by PCA, PRC, species numbers or abundances of individual species, no correlation could be found, not even trends were visible.

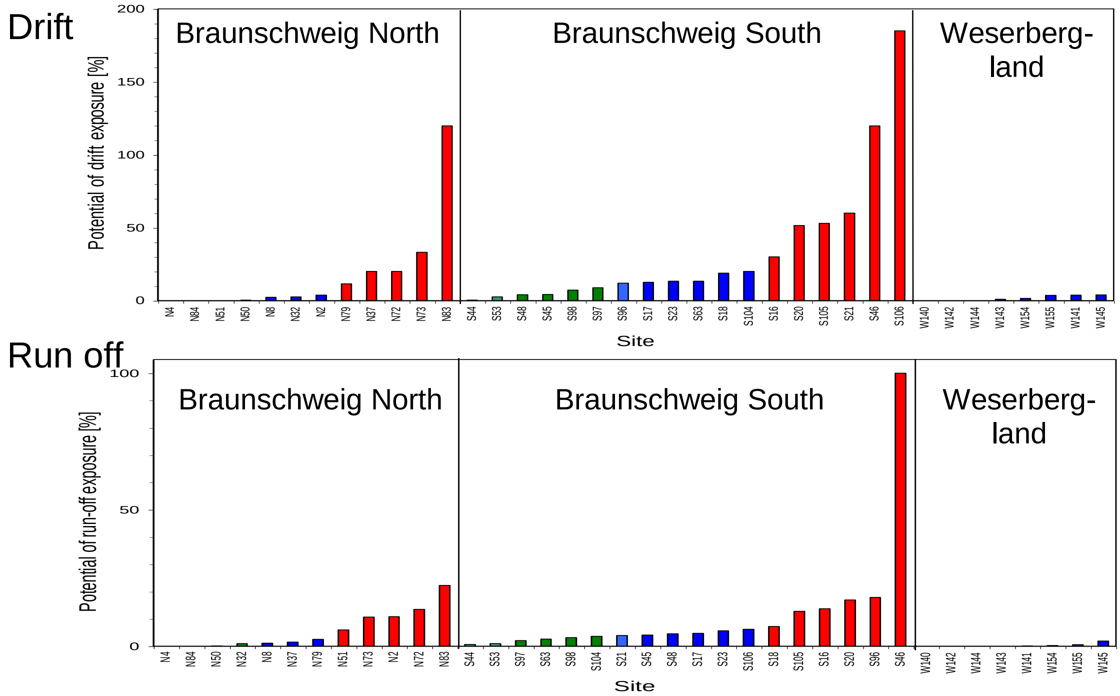


Figure 48: Range of estimated potential of exposure by spraydrift or runoff and grouping for further analyses
No exposed sites were available for the Weser mountain subregion.

A clear shortcoming of this part of the project was the calculation of the potential for exposure. It had to consider a much higher number of influence factors (structural differences, water flow, upstream use) than that calculated for the “Altes Land”. The basis information was less exact. At the same time, the use patterns were more variable and the applications were not synchronized. This might be partly improved by including data from a project performed by Syngenta, which developed exposure analyses based on geographical information systems of high resolution in the region of Braunschweig for three years.

3.4.4 Discussion and implications for futher work

Differences in macroinvertebrate community structure are linked to a range of environmental variables. Due to the unique situation of the sampling area “Altes Land”, most of the habitat variables typically affecting freshwater macroinvertebrate communities could be neglected: Especially topographical variables are very similar for the different sites. In addition, velocity (nearly zero) and the type of sediment were the same for all sites. However, land use (grassland or orchard) was found to be a very important variable: Sites along grassland showed different macroinvertebrate communities compared to sites along orchards, and explained more variability in community structure than the potential for exposure. Reasons for the special situation of the grassland ditches might be a minor shading, which might increase light intensity, water temperature and, consequently, primary production. The likely lower input of leaves into the grassland ditches might be one explanation for the tendency towards lower nutrient concentrations measured there, next to a potentially lower input of fertilizers in grassland as compared to the orchards. Macrophyte densities did not seem to be related to grassland or orchards along the banks, but the species composition was different. The different habitat structure close to the ditch might also be important for adults of insect species with aquatic larval stages.

Focusing on the orchard sites, the estimated potential of exposure was identified as the dominant gradient driving the community structure. The PRCs demonstrated the relation between potential of exposure and community structure over the seasons and identified the taxa that contributed most to the observed changes in the community structure. Clear and permanent effects on the communities could be demonstrated in the ditches with a high potential for pesticide exposure, which were very close to the trees (≤ 1.5 m). Differences between the sites with medium exposure potential (distance to the trees of 3 – 5 m) and the sites with a low (or negligible) exposure potential were generally smaller. The higher abundances of Pulmonata in the ditches with a high exposure potential might be explained by the usually low sensitivity of mollusks towards insecticides and the decrease of competition of more sensitive macroinvertebrates affected by the likely significant insecticide inputs. The fact that Bivalvia were only found in ditches with a low or medium exposure potential, but not at potentially highly exposed sites, may be explained by enhanced susceptibility of early life stages and/or a lower recovery potential. Despite seasonal trends of recovery, shifts in species occurrence correlated to the theoretical exposure potential were exhibited permanently at highly exposed sites.

Representativity of the results

Macroinvertebrate communities of ditches and streams in intensively cultivated agricultural landscapes, especially when used as drainage systems, are expected to be able to cope with disturbances and/or have a potential for rapid recovery. The ten most dominating species determined in total are widespread and common in many types of water bodies in Northern Germany.

However, species with susceptible life stages and slow population dynamics, as represented in the lists of endangered and vulnerable species were also found. These species have a limited recovery and recolonization potential. Their presence in the investigated agricultural ditches suggests that pesticide impact on ditches with a mean distance of 3 m or more to the trees is limited. They also indicate a considerable stability of these ditches with respect to species composition. The findings demonstrate that the “Altes Land” region provides habitats required by these species and thus valuable resources of biodiversity within an intensively used agricultural landscape.

The risk for potential pesticide exposure in the ditches of the region "Altes Land" can be considered to represent a worst-case situation due to the frequency of application, the horizontal spray direction and the small distance to the surface waters. In 2000, the Federal Law of Lower Saxony adopted special regulations, which allow a reduced distance of 5 m between the application area and permanent water bodies for listed pesticides, if additional mitigation measures are used such as drift reducing techniques. This distance corresponds to the upper range of the distances at the sites characterized by a medium exposure potential in the present study. Thus, it can be expected that the potential for effects is lower than observed. From 2001 to 2003, the Federal Biological Research Centre for Agriculture and Forestry (BBA) conducted a chemical and ecological monitoring at four sites in the region "Altes Land". Three of these sites corresponded to sites, which were analyzed here. Only at the site I-5 with a distance below 1.5 m between the trees and the ditch, the determined pesticide concentrations indicated a risk for the aquatic populations (the sum of toxic units based on the lowest NOEC for each active ingredient detected exceeded a value of 1). In principal, the findings of the BBA confirm the interpretation made here.

From the results obtained for the orchard region of the "Altes Land“, it can be concluded that a relatively small buffer zone of 3 to 5 meters would likely be protective also for macroinvertebrate communities in ditches or ponds of other agricultural regions with lower exposure potential. This may be confirmed by the results of the Braunschweig project part, where no indications for landscape level effects of pesticides could be found.

Conclusions

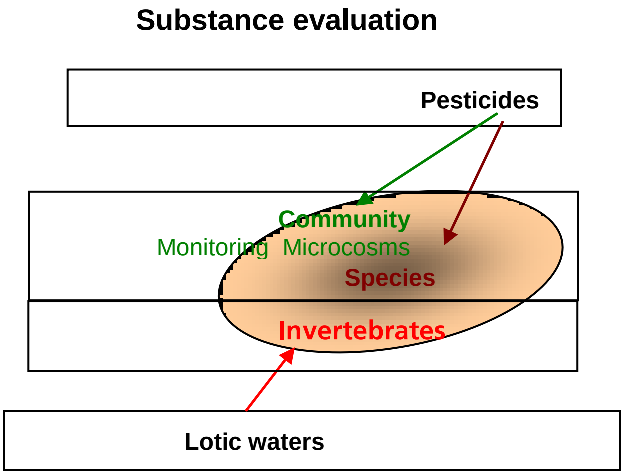
The analysis of aquatic invertebrate communities in the investigated agricultural scenario detected significant effects of pesticide use, trends of recovery and long-term differences between site groups, demonstrating the practicability, sensitivity and reliability of the selected approach. However, the situation in the Altes Land is unique because the water bodies are similar with respect to stream morphology, water and substrate quality as well as topography (compared to small water bodies in larger, more diverse landscapes). At the same time, the land use and exposure situation is very homogenous and can be regarded as „worst case“ due the high use of pesticides in orchards, the “Altes Land” thus being an ideal example for linking cause and effect. The community structure in the ditches located closer to the trees was significantly different due to the loss of species and shift in abundances. Our results show that a distance of 5 m between ditches and orchard trees should prevent significant long-lasting effects on the macroinvertebrate community in the “Altes Land”, which is characterized by very intensive orchard cultivation.

For the water bodies in the region of Braunschweig the situation is more complex: the water bodies are more diverse according to structural parameters and the exposure situation is less easy to estimate. Thus, it is strongly recommended to focus on structurally similar water bodies and to improve the estimation of exposure potential in further monitoring studies.

The findings in the Braunschweig region are partly in contradiction to the various publications by Liess and co-workers (e.g., Liess and Schulz 1999, Liess et al. 2001). It has to be emphasized that the scope and approach is different. While Liess tries to find specific effects of pesticide peak concentrations on specifically sensitive indicator organisms in individual, specifically investigated streams, the presented approach tries to find statistical evidence for hazards of pesticide exposure on the landscape level.

3.5 Artificial stream system for investigating habitat-specific fate and community effects

3.5.1 General scope and approach



In the registration process of pesticides, the standard aquatic toxicity data of active substances (a.s.) are generated at maintained concentrations in water to investigate intrinsic properties of the a.s. and to calculate toxicity : exposure ratios (TER). If the predicted environmental concentration (PEC) of a pesticide use resulted in TERs beneath the trigger values as given in the Council directive 91/414/EEC, a detailed aquatic risk assessment should be performed that includes more realistic considerations of fate and effects. A typical contamination of a waterbody with a pesticide is characterized by the temporary and spatially limited application and the removal of the active substance(s) from the water phase by adsorption at the sediment and degradation. If a higher tier study is necessary to demonstrate that adverse aquatic side effects of the intended use can be excluded, a combined investigation of fate and community effects under natural conditions is the method of choice. This is normally approached by freshwater lentic field tests, such as mesocosm studies. A possibility of an indoor test system with controlled lentic conditions is described in chapter 2.4.3. For a better approximation to reality, experimental test systems simulating natural stream ecosystems may be of considerable value (Crossland et al. 1991), especially for pesticides with insecticidal and adsorptive properties. A realistic simulation of pesticide loading should be based on an experimental system having sufficient size and complexity for self-preservation of the inherent populations. Sediment composition and structure should be natural and inhabited by naturally age-structured populations. Investigations of fate and effects of the test substance(s) should be possible in different distances from the loading area, at sites with different stream velocity and in different microhabitats. Additionally, the system should be able to investigate water leaching through the riverbed to simulate bank filtration processes and potential groundwater contamination. A realistic fate assessment should be based on studies at relevant concentrations, which may he very low, especially when the pesticide is highly toxic. At the same time, metabolite formation may be of interest. Thus, the use of radiolabeled substances is recommended. In the early 1990s, an artificial stream system was constructed at the Fraunhofer Institute in Schmallenberg and patented. The possibilities of experimental set-ups should be explored by a postdoc for one year.

This was my entry in Fraunhofer. At first, the conditions of maintaining a complex stream macroinvertebrate community were investigated (ref. 10). As the freshwater field test set-ups were intensively discussed in the early 1990s, it early became obvious that the artificial stream system as it is could not serve as stream mesocosm in the regulatory assessment of ecotoxicological effects, as at least 10 replicate streams would have been necessary. However, it might be a valuable tool for experimental fate simulations and as reference system to validate hypotheses on effects on stream organisms under natural conditions.

Thus, the first series of treatments was performed focusing on fate, but including site- and habitat-specific aspects (ref. 13), 35), 36)).

3.5.2 Artificial stream facility

The artificial stream mesocosm (Figure 49) was installed as part of the institute's outdoor field simulation facilities, permitted as controlled area for the use of radiolabeled substances. The needs to meet the requirements of radiation protection led to a loss-free recirculation design completely made of stainless steel. It is placed in a concrete basin, which can be drained under control. The total length of the oval stainless steel trough is about 35 m with a width of 0.6 m and a depth of 0.35 m. It contains an extended segment and two inlets to diversify the current. The inlets are equipped with one aquatic lysimeter each, flanged at the bottom of the trough. The lysimeters can be filled with undisturbed sediment monoliths of 0.6 m diameter width and 0.6 m column height. Water leaching through the sediment cores is sampled in separate stainless steel containers, the flow-through is regulated by valves. An electrically driven double chain with steel paddles is responsible for the generation of water flow. Steel covers prevent splash water. Stream velocity varies between 0 and 0.3 m/s with a mean at about 0.05 to 0.1 m/s. On the bottom of the concrete basin, two stainless steel tanks of 5000 L resp. 2500 L volume were installed. The bigger tank collects overflow caused by heavy rainfalls and receives the contaminated water at the end of the studies, whereas the smaller one is used as reservoir for uncontaminated freshwater to compensate evaporation effects. Both tanks are aerated. The concrete basin is filled with a 1 m drainage gravel layer and agricultural soil up to the upper margin of the trough. The facility is covered by a glass roof which can he opened und closed automatically via rain and wind sensors.

Water and sediment from a natural brook including the original invertebrate community, are used to fill the artificial stream. The brook is located within or directly downstream a water protection area (Wenne, Schmallenberg) where pesticide uses are not permitted.

Figure 49: Artificial stream mesocosm in state of construction

Sampling and analysis. For the investigation of sediment samples, stainless steel baskets (9 cm diameter, 15 cm height) filled with sediment were placed in the sediment layer before the beginning of the test. The number of baskets at each sampling site is determined by the sampling frequency. For sampling two cylinders are placed around a basket and the water within is removed by a pump. The inner cylinder is connected to the ground plate of the basket and prevents erosion of the sediment. It is taken out of the sediment layer together with the basket by pulling the handle of the basket. In the remaining outer cylinder, a new basket filled with sediment can be introduced without disturbing the direct environment. The removed sediment can be separated in different layers and analyzed for substances (fate) and/or organisms (effects). This method is also suited for the determination of recolonization rates. Samples of water and periphyton are taken from the same sampling sites.

In the case of fate studies, sediment and water samples are investigated with respect to total radioactivity, active ingredient(s) and main degradation products. The amount of total radioactivity in water samples is analyzed by liquid scintillation counting, while sediment and periphyton samples are analyzed by combustion. For the identification and quantification of test substance and metabolites, the samples are extracted by liquid/liquid extraction, solid-phase extraction or concentrated by lyophilization. The organic extracts are analyzed by appropriate chromatographic methods.

To gain realistic data on fish bioaccumulation, fish cages with e.g. juvenile rainbow trout (Oncorhynchus mykiss) can be installed in the artificial stream mesocosm in different distances from the loading site. Samples of three fish per cage are taken at each sampling date. They are blotted dry to obtain the fresh weight before drying at 80°C overnight, followed by the analysis for accumulated total radioactivity by combustion.

Advantages. The artificial stream mesocosm is suited for studies on the fate of radiolabeled substances in the stream compartments water body, sediment, leaching water and organisms such as periphyton, macroinvertebrates and fish. Because of the natural communities, studies of the realistic effects on meiobenthos and macroinvertebrates can be combined with the fate studies. Undisturbed parts of the natural brook where the sediment is taken from may serve as untreated control site. Test conditions closely simulate natural conditions concerning rate, bioaccumulation and effects of the tested pesticide, even with respect to zones with different stream velocity, particle sizes and distances from a spot loading. With respect to the objectives of a detailed and realistic risk assessment, especially when the pesticide is adsorptive, persistent, an insecticide or used in regions with streams as dominant waters, the mesocosm should represent a realistic option.

Drawbacks. Initial experiences with the artificial stream (ref. 10); diploma thesis by Angela Ertz 1996) and a first study with a pesticide revealed some weaknesses of the system.

- The top of the hill site of the facility is responsible for extreme conditions of the microclimate and insulation, which are not typical for connected valley sites of streams. Geographical isolation and wind effects make it difficult for aquatic insects to (re-)colonize the artificial stream system. During summer high irradiation at the water surface caused high water temperatures.
- The air temperatures during winter resulted in a nearly frozen system. Ice had to be removed from the paddle drive daily. Snow melting and heavy rainfalls resulted several links in flooding of the natural control brook.
- In the beginning, the area within and around the artificial stream was not planted. Thus, there was no natural shadow, no substrate for adult aquatic insects and no natural leaves input for the nutrition of the system community.
- The recirculation design was a concession to the demands of radiation protection and water supply, while linear transport is a characteristic of streams. In a first fate study, the real dilution of the substance depending on the distance from the loading point/site could not be simulated correctly.
- The diversification of stream velocity was not yet sufficient to simulate major local differences in the particle size fraction of the sediment, influencing pesticide sorption and fate.
- The size and form of the fish cages induced sack water, which impaired the distribution and fate of the test substance in the system.

Improvements. Some of the drawbacks led to immediate improvements, but others are of principle nature.

- Effect studies still have to be evaluated cautiously when comprising the time of oviposition and colonization of insect larvae from May to September. The high irradiation is limited by shading.
- An electric heating was installed to prevent freezing of the water. An identical control system made of cheaper material is planned outside the control area for radiolabeled substances.
- The area within und around the artificial stream was planted with meadow plants. Black elders and willows in pots and camouflage nets simulate natural shading. If the trees are too small to produce enough leaves for the nutrition of the system community, certain amounts of leaves will be supplemented.
- The main efforts have been focussed on a technical solution for the option on an initial flow-through design. If the water flow is started after loading and the contaminated water is replaced after one circulation by untreated natural water, a simulation of linear transport, a natural dilution and a real spot loading is possible. For this purpose, the existing and two additional tanks of stainless steel are used (Figure 50). Temporarily, a blocking plate und a surface barrier prevent contaminated and untreated water from mixing, while it is led into the big tank. Thus, a once-through design can be maintained until the first 5000 L of contaminated water (about two fillings of the artificial stream) are discharged into the tag tank. The site within the barriers could he used as internal control site upstream the loading site. After two circulations the tank has to be closed, the two barriers have to be removed and the recirculation status is re-established. The possibility of a slight contamination of the whole system by desorbed test substance and more hydrophilic degradation products depends on the properties of the test substance. For the first experimental set it could he shown that when the exchange of all 5000 L was realized, the remaining test substance in water was negligible. Nevertheless, it would only simulate a natural background loading by formerly contaminated water.
- Zones of higher stream velocity (up to 0.5 m/s) will be created by additional propulsion.
- The first cages were given a more streamlined design (Figure 50).

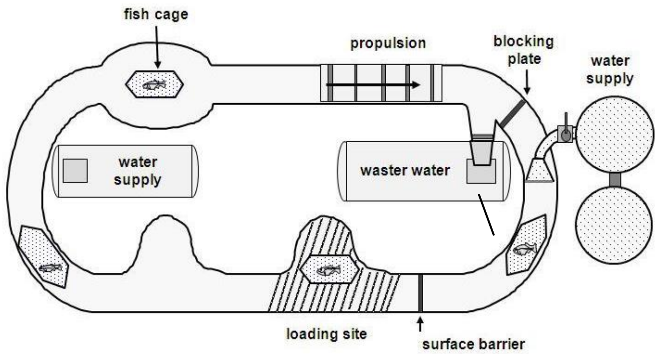


Figure 50: Sketch of the new artificial stream design to simulate spot loading

Comparison with the natural reference stream. Comprehensive investigations of the survival of the macroinvertebrate populations in the artificial stream and comparative studies on the development of diversity in both the artificial and the natural reference stream were performed (ref. 10) and diploma thesis by Angela Ertz). The artificial stream facility was filled in autumn with water and sediment including organisms from the natural reference stream in autumn. After an adaptation period of 1-2 months, macroinvertebrate species diversity and population densities were similar with the exception of amphipods (lack of leaves in the artificial stream). After 5 months of hibernation, the communities were yet similar, but diverging during the period of emergence of the aquatic insects the next summer. The results demonstrate the natural conditions of the artificial stream as well as some weaknesses of the system as mentioned in the drawbacks section.

3.5.3 Pilot study on the fate of a pesticide

A first study with a pesticide was conducted to show strengths and weaknesses of the artificial stream test system, especially whether it is suited for a simulation of a real spot loading and a different distribution of the test substance dependent on stream velocity and dynamics (ref. 36).

During the study, the amount of radioactive material in the sediment increased as expected. At the end of the study (50 days after application), approximately 30% of the total radioactive material were found in the upper sediment layer (0-5 cm). Compared to the results obtained from the water phase, the upper sediment layer contributed most to the rapid degradation of the test substance in the whole system. Only slight variations in the concentration of total radioactive material at different sediment sites were observed, but they could be explained by the site-specific exchange rates between water and sediment, being dependent on stream dynamics and velocity and due to phenomena like slack water and existing surface layers. Shortly after application, the highest concentrations of the parent compound were found in the upper sediment layer at turbulent sites near the application area, whereas at sites with stagnant water and sedimentation the concentrations were lowest due to reduced water exchange and contact probability. At neighboring sites extremely different in stream dynamics (turbulent site at a fish cage and sedimentation site 30 cm aside) the concentrations of the parent compound in the upper sediment differed by more than an order of magnitude one day after application. Additionally there was a clearly fading gradient of the sediment concentration downstream. In the periphyton, the highest total radioactivity occurred 24 hours after application and the relative values at different sampling sites were similar to the distribution of the total radioactive material in the sediment at the same time. In the course of the study, the concentration of radioactive material in the periphyton decreased faster than in the water, due to elimination of more polar metabolites. In the further course of the study, radioactivity, mostly representing the polar metabolites, significantly distributed to the lower layer of the sediment (5-15 cm) and exhibited the highest concentration at the site with highest stream velocity and most laminar flow (prior to the paddle wheel propulsion). The great amount of more polar metabolites was also responsible for the fast elimination and low accumulation in fish, compared with laboratory data. Thus, the maximum of total radioactivity in fish was measured 48 hours after application.

Effect on meiobenthos. At the time of exposure, the macroinvertebrate community was already deteriorated by emergence of aquatic insects and lacking recolonization (except chironomids). We thus investigated the zoobenthos community associated with the periphyton layer on the sampling baskets. The main groups identified were chironomid larvae, copepods, nematodes and bryozoans. The abundances at each sampling date were related to the taxa-specific abundances at the same site during the pre-application phase. Two trends were observed which partly interfered. The losses of the less mobile organisms (nematodes and bryozoans) were roughly correlated with the concentrations of the parent compound in the upper sediment and periphyton 24 h after application. During the study, these losses did not recover, as due to the low temperatures in autumn there was no repopulation. For the big and mobile taxa (copepods and chironomids), abundances decreased at sites with high concentrations but increased at the sites of medium and low concentrations downstream, clearly indicating flight reactions or drift of individuals that were partly able to actively attach to the downstream sites. During the course of the study, the big taxa recolonized the sites near the area of application again, leveling out the differences in density at the different sampling sites.

3.5.4 Conclusions for further work

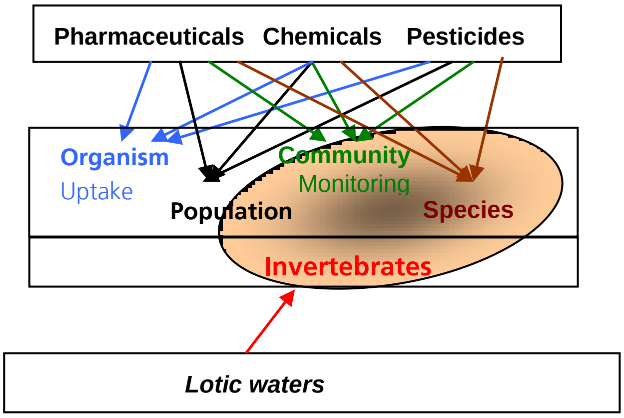
The first results demonstrate the capability of the system to simulate the property-specific temporal and spatial exposure pattern of substances being dependent on stream dynamics, sediment properties, and distance to the site of contamination. Whereas this seems to be of limited interest for predictive substance evaluation (modeling is obviously regarded sufficiently safe), it can be of high interest for the specification of micro-sites for sediment monitoring as well as for research on interactions between substance properties and site- and habitat-specific organisms and communities. The sediment surface of the artificial stream system can be structured and the water propulsion be directed to create as different conditions as feasible at the same time. By using different labels, different substances and their metabolites can be investigated at the same time.

Micro-sites for sediment monitoring. Within aquatic monitoring programs, sediment-sampling sites at streams are mostly chosen according to the situation of a source of contamination (upstream or downstream) and accessibility. Generally, there is no systematic record of the local stream dynamics or habitat properties. However, as concentrations most probably differ hugely depending on physico-chemical properties in a small spatial context, there is need for a systematic investigation of substances representative of a wide range of physico-chemical properties concerning the fate in stream systems. The aim is to generate a checklist for crucial parameters of flow, turbulence or sediment surface structure, which should be obligatory for sampling occasions. Using the parameter values, the findings for classified relevant physico-chemical properties can be normalized. By this approach, the monitoring of sediment concentrations is no longer a matter of stochastic coincidence.

Exposure and sensitivity of different stream habitats. It could be demonstrated that after sufficient equilibration time the species characteristic of different stream habitats tended to organize in the same habitats simulated by the artificial stream system (ref. 10). As specific stream habitats are characterized by flow, turbulence or sediment surface structure, exposure to chemicals is different due to the physico-chemical properties. At the same time, the species assembly is habitat-specific. It is evident that organisms of fast-flow habitats need a higher oxygen concentration and thus a higher water exchange rate at respiratory membranes (see chapter 2.2.3). In consequence, it has to be clarified whether habitat-specific species traits and habitat-specific exposure potentially create an extraordinary risk, or whether a specific constellation tends to reduce uptake, accumulation and thus internal exposure. If habitat-specific trends in population dynamics can be identified (e.g. related to size, mobility, taxonomic representation), also species traits determining recovery have to be included in the considerations.

3.6 EU 6th framework proposal: Water quality evaluation concept based on a habitat approach

The whole project should focus on dynamic surface waters, starting with spring regions, persecuting lotic waters from upstream to downstream regions, including estuarine and ending with coastal marine waters. Lentic waters outside a river network should not be included since aspects of the aquatic community and pollution situation of those may be very specific.



Water quality

3.6.1 The habitat concept

The habitat is the smallest environmental unit with respect to defined structural and micro-climatic conditions, for which a typical biological community can be defined. In aquatic systems, it is mainly characterized by surface structure and flow regimen and by the inhabiting community itself, especially primary producers, and thus light conditions.

The communities of comparable habitats may differ due to water quality, nutrient state and temperature, as well as due to the species colonization patterns: They may be different in different river basins and water bodies.

The observation of a water body is performed at representative locations (sites). A water body consists of several habitats. A representative site comprises all characteristic habitats. Thus, for measuring ecosystem health of water bodies, a site-specific investigation integrating all habitats is demanded by the WFD and general practice in quality surveys. However, a generic assessment of specific impact should focus on the individual habitats for the following reasons:

- The concentration of pollutants in the sediment and on surfaces is dependent on the habitat-specific surface structure and flow regimen (speed, turbulence), driving the contact rate of dissolved, dispersed or transported materials with the surface, as well as on the physico-chemical properties of the pollutant. This might result in concentrations in sediment varying by more than an order of magnitude at different habitats of the same site (chapter 3.5 (ref. 13)).
- At the same time, the species characteristic for a community of a specific habitat have specific needs being correlated with surface structure and flow regimen, finally driving the water exchange probability at respiratory membranes. Thus, habitat specific properties in combination with the mode of feeding are responsible for the uptake rate of pollutants (chapter 2.2.3; (ref..65)).

The concept of habitat specific fate and effects should be translated to and specified by a combined model (first step: fate and bioavailability model; to be combined with a MoA-specific effect model) and used as a generic tool for the ERA of chemical pollutants for all relevant water bodies.

Relation to existing work: The biological ecosystem health surveys are already performed in a habitat-specific way and have only to be reported habitat by habitat, or the data can be attributed to the habitat by using key species. For example, the Bavarian water authority published habitat preferences for most known German species.

The concept may connect the actually completely different worlds of nature conservation (habitat directive) and substance specific ERA.

The presented project concept (Figure 51) was worked out by Caroline Peters, TU Berlin (now: TU Hamburg) (specific work packages: 1 Coordination, WP 3, WP 9), and Christoph Schäfers (WP 2, WP 4), and further completed by other partners.

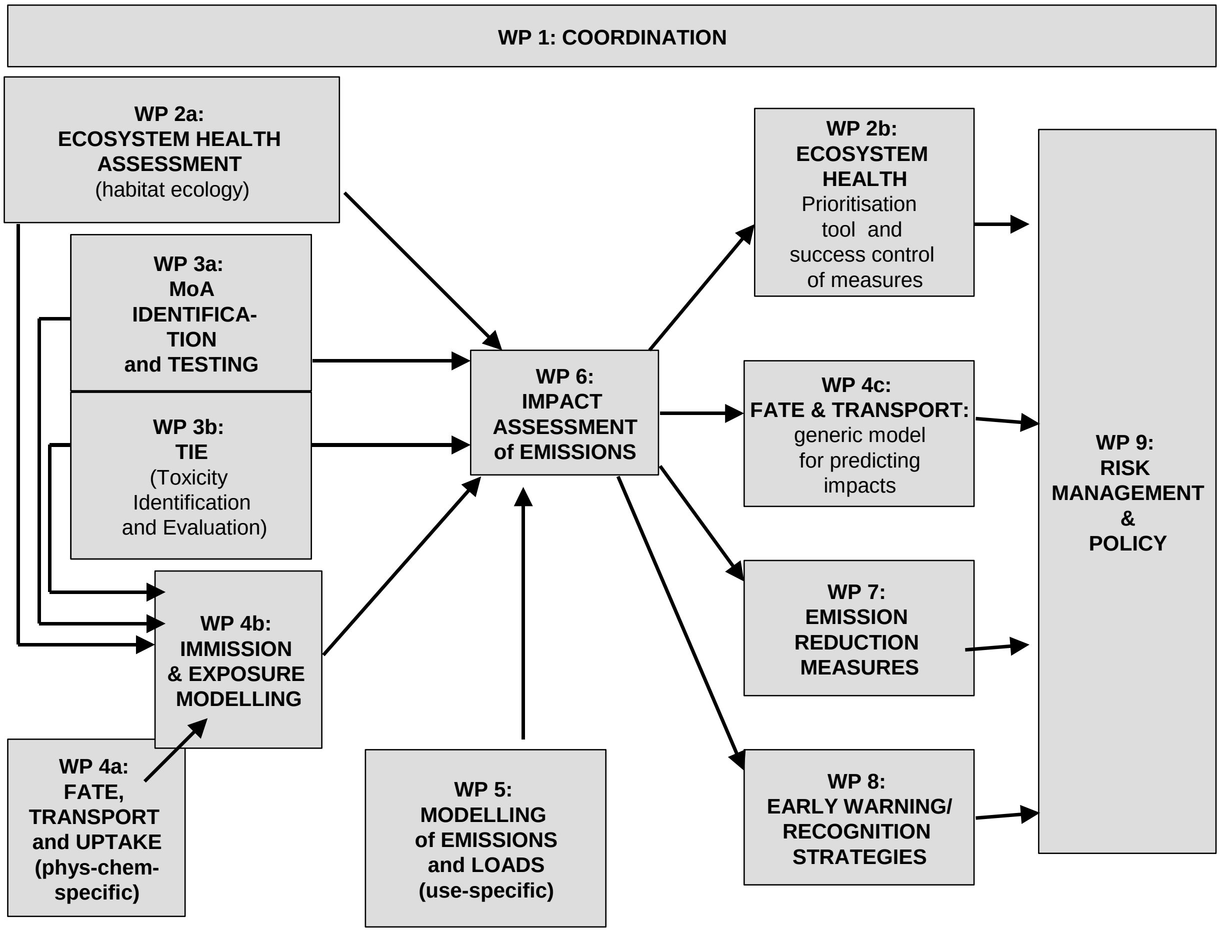


Figure 51: EU 6th framework proposal: Project concept and structure

3.6.2 WP 2: Ecosystem Health

This work package focuses on methods for the assessment of Ecosystem Health according to the WFD. It has to include benthic macroinvertebrates, macrophytes and fish. Macroinvertebrates and macrophytes (this might also be extended to periphyton), can be used to characterize sites and attributed to specific habitats. Fish are mostly more mobile, using specific habitats and sites including these habitats mostly during specific life-stages only. The overall objective of WP 2 is to evaluate ecosystem health with respect to community and population structures and not with respect to specific indicators of toxicity or bioaccumulation. Developing toxicity indicators and uptake models is the objective of WP 3 and 4.

Needs for WP 2: partners with experience in biological monitoring in different parts of Europe, also in marine environments. Helpful: Further quantitative ecologists.

2a: Ecosystem Health assessment

Aquatic communities are indicators of Ecosystem Health as well as its subject. They integrate influences of all kinds and over time of importance for the organismic integration as well as for the species-specific population dynamics. For Ecosystem Health assessment, there are two main principles in use: The stress indicator principle and the reference deviation principle.

The stress indicator principle uses stress-specific indicators either of high specificity, which is very precise but needs much knowledge on modes of action and excludes any influence being out of the focus, or of low specificity, which does hardly allow any causal analysis and may indicate an influence, which is not of any interest with respect to Ecosystem Health. Examples for this principle are the different biotic indices indicating organic pollution (British BMWP, Belgian Biotic Index, German Saprobienindex), trophic status, or acidification. With respect to specific chemical pollutants, they are often very indirect if fitting at all.

The reference deviation principle uses holistic information on the unpolluted reference status (defined as “healthy”) and identifies deviations from it. It aspires to include all relevant influences without identifying them and can help to prioritize measures with respect to the need for improving ecosystem health. Work according to that principle is comparably complex. An example for this principle is the British Riverine InVertebrate Prediction And Classification System (RIVPACS). With respect to specific chemical pollutants, it provides an overall Ecosystem Health status, but needs identification of the relevant stressors.

Benthic macroinvertebrates and macrophytes (perhaps also periphyton):

In this project, a RIVPACS-like method should be used to classify the ecosystem health status. During a first step, different habitats will have to be chosen for which ecosystem health will be evaluated. The habitats should represent all important habitat types in dynamic water bodies. In a second step, regions/water basins all over Europe should be chosen for exemplary investigations. Criteria for the choice should be

- representation of habitats,
- accessibility for a project member; availability of regional geographic and taxonomic knowledge,
- representation of pollution sources,
- high variability in the pollution levels, ranging from pristine to strongly polluted sites.

For each habitat type, pristine or very marginal polluted sites will have to be identified. Because a multivariate approach is intended, a small number of river basins (not more than 4), a medium number of habitat types (e.g. 15), and a high number of sites with different types and levels of pollution (at least 200) should be investigated. Each work in a different river basin may focus on different classes of pollutants. For WP 2, this may result in river basin-specific selection criteria of sampling sites. The selection of sampling sites as well as of influence parameters to be included in the measurements and the harmonization of sampling techniques and taxonomic determination will take at least 1.5 years.

The sampling should comprise an early spring, midsummer and autumn sampling over two years of organisms and simultaneous measurements of potentially influencing parameters.

A TWINSPAN analysis on all aquatic communities of the different habitats should differentiate groups of communities by

a) habitat

b) river basin

c) type and grade of pollution

For each habitat (and most probably river basin), a group of reference communities should become obvious. The deviation of the habitats of other than reference sites from these should be quantified. This work will take 2.5 years.

The deviations that cannot be explained with geographic, hydrological or structural influence factors (and thus could be seen as sub-groups of the intended habitats), can be correlated with pollution-related influence factors. The data will have been generated by the other WPs, either by generic models or by monitoring the same sites. The correlations and impact factor identifications can be part of WP 6. (Fh-IME may contribute to the selection of sites and parameters as well as to the TWINSPAN analyses and multivariate causal analyses)

Fish

Because of the mobility and the generation time, habitats and endpoints will have to be different from macroinvertebrates. It is proposed to monitor the developmental stages of fish species in different seasons, and to compare reference sites with polluted sites (worked out by a project partner).

2 b: Prioritization tool / success control

The results of the causal analysis can be used to generalize the method (e.g. to simplify for a focused objective) to develop a site prioritization tool for pollution identification and management measures. At the same time, this tool can be used as success control tool for performed measures.

3.6.3 WP 4: Modeling of transport, fate and uptake

This work package focuses on models for the assessment of fate of and internal exposure to chemicals focused on. It has to include all relevant aspects of fate (i.e., transport and distribution, degradation in water and sediment) and uptake by organisms (i.e., bioavailability[7], water exchange rates at respiratory membranes, uptake rates) with respect to sites located in different distances from the sources of pollution and differentiated into specific habitats.

The overall objective of WP 4 is to combine models and to calculate generic immission concentrations for different habitat types at specific sites (to be confirmed by chemical monitoring) and potentials for internal exposure of habitat-specific organisms as pre-requisite for toxic action. Necessary input has to be given by WP 5 (i.e., emission data for the relevant chemicals from relevant sources). WP 6 can use the output of WP 4 by combining internal exposures (WP 4) with the identified MoAs (WP 3) to assess habitat-specific effects, and to compare these with ecosystem health assessment of the respective habitats (WP 2).

Needs for WP 4: partners with experience in modeling the different aspects, generating missing data for focusing the models, and chemical monitoring.

4a: Fate, transport and immission modeling

Impact on aquatic biota is a function of exposure and toxic action. Focusing on the habitat scale requires habitat-specific views on both aspects. WP 4a focuses on exposure to the

---

[7] For some metals, biotic ligand models (BLM) were developed. A comprehensive summary of concepts, models and applications for copper and zinc is given in (25).

chemicals in question, considering concentrations in the media relevant for habitat-specific organisms.

The first task is to select suited models of fate and transport and to select relevant physical-chemical properties or degradation data suited for the use in models of fate and transport.

The second task is to collect or generate respective data for the substances/substance classes we will focus on.

The third task is to collect or generate data about parameters that are describing specific habitats, requested by the models for calculating habitat-specific immission concentrations.

4b: Uptake modeling

WP 4b focuses on exposure to the chemicals in question, looked at as doses entering the habitat-specific organisms.

The first task is to review models of uptake or of substance exchanges at respiratory membranes, to select relevant physical-chemical properties suited for the use in uptake models and to generate a suited model being applicable for different types of organisms

The second task is to collect or generate respective data for the substances/substance classes we will focus on.

The third task is to collect or generate data about parameters that are describing specific uptake characteristics, requested by the models for calculating doses for habitat-specific species. This may be generic (by including knowledge on physiological properties), but should be confirmed by integrative experiments (^{14}C-labelled uptake experiments with representative species).

4c: Generic model

WP4 combines emission/mass-flux models (WP5), fate and transport models, and uptake models for calculations of habitat-specific generic internal exposure (dose). As a validation experiment, ^{14}C-marked substances can be investigated in our 35m artificial stream system with different habitats for immission concentrations in the different habitats and doses in habitat-specific organisms. The results may be used to recalibrate the combined model.

The internal exposure model can be used as generic tool for bioavailable exposure assessment of chemicals based on the base set of information on the chemical (i.e., physico-chemical properties) and the relevant information on the site and habitat in question (i.e., distance from sources of pollution, relevant habitat properties as derived in WP 4).

3.6.4 Implications for further work

The funding of the project was narrowly missed. However, especially the issues mentioned in task 4b identify potential for extensive experimental research, which may be organized in many small individual projects.

4 General Conclusions

Relevant exposure. After years of separated exposure and effect considerations in environmental risk assessments, there is need to focus on exposure relevant for causing effects. This especially applies to the uptake processes in organisms.

- Biovailability is a main key for understanding effects (i.e. metal availability, chapter 2.4.5).
- For soluble substances, water exchange rates at respiratory membranes seem to be an important factor for uptake and thus internal exposure of target structures (i.e. insecticide toxicity, chapter 2.2.3). In this respect, bioconcentration is more important than biomagnification when considering effects.
- Comparative toxicology needs the relation of effects to body doses or has to include information about uptake and elimination rates.
- For particulate matter, including nanoparticles and metal agglomerates of a diameter < 0.1 µm, ingestion via drinking is of high importance and may explain e.g. phenomena of metal accumulation (see outlook chapter 5).

Physiological effects. The universe of physiological variability in different taxa has to be taken into account. At the same time, a sensible generalization is necessary to focus on principles of relevant effects.

- In case of concern, specific modes of action need specific attention when planning and evaluating studies for hazard assessment. Targeted approaches (e.g.) have to be set in ecological contexts, and based onappropriate population relevant endpoints (i.e. endocrine disruption test an assessment strategies in fish, chapter 2.3).

Behavioral effects. Behavioral reactions and interactions should be accounted for when effects on populations in the field are to be predicted or interpreted.

- Flight reactions depend on the mobility as well as on the sensory capability of species or life stages (i.e. copepods, chapter 2.4.4.2; epibenthic communities, chapter 3.5.3).
- Reproductive toxicity including endocrine mediated effects often is poorly understood unless interactions between sexual partners, e.g., mating and spawning, are included (i.e. in fish life cycle tests, chapter 2.3.6).

Population and community effects. With increasing level of biological system organization (molecular, cellular, organic, organismic, population, community) the complexity rises with functional redundancies and thus the capacity increases to compensate effects. The compensation capacity ends with the deterioration of accessible resources for running functions (e.g. gene pools, habitat capacities in terms of structure and feed).

- At higher levels of systemic organization, concentration-effect-relationship become steeper as long as common regulative properties of the integrated levels are affected, which is true at least for the population level. Basing chemical risk assessments on higher levels of biological organization enhances the probablilty of including the most sensitive physiological interactions. However, NOEC values increase due to masking of physiological or individual effects by compensation, besides the fact of less powerful statistics due to the mor difficult data akquisition.
- Populations can be affected by age class and sex structure, size, and (sub-)population survival time and rate.
- A steeper concentration-effect relationship increases the probability of severe effects on (sub-)populations at toxicant concentrations close to the observed effect threshold level.

- To evaluate ecotoxicological influences in the field, the whole ecology has to be accounted for to identify and confounding factors and differentiate ecotoxiclogical from other effects.

Structure versus function. The central function of an ecosystem is to provide biodiversity. This function can most validly be evaluated by investigating the biological structure (i.e. community level studies, 2.4). In the future, observations including population genetics may enhance ecological realism by identifying effects on the compensation potential for future hazards (see chapter 5).

Recommended approach for substance evaluation. In a limited world with a growing human population and the challenge of a sustainable development of technical progress and management of resources, substance evaluations based only on hazard assessments often are not sufficiently realistic to accept uses of the substance. To achieve a pragmatic as well as protective regulation, substance evaluation should be based on risk assessment. Risk assessment is much more complex and should be open for a weight of evidence approach aiming at a maximum of consistence. Especially in situations where benefits of environmental concentrations are opposed to the risks (e.g., for pesticides and specific biocide uses or consequences of the use of building materials), the risks have to be assessed thoroughly:

- The investigated test species should be relevant in terms of physiological sensitivity and habitat exposure (see chapter 2.3). The uncertainty of the species sensitivity distribution should be quantified and evidence generated that there is no effect below a threshold concentration.
- Particularly sensitive life stages or performances have to be identified and included in the risk assessment (see chapter 2.2).
- If there is too much uncertainty left, a community level study should be conducted under realistic exposure conditions in the context of a food web comprising representatives of the most sensitive organisms and life stages. When exposure is of a temporal nature (as pesticide uses), the time to full recovery can be determined and evaluated whether it is sufficiently short to exclude initial effects being adverse (see chapter 2.4).
- The assessment can be proven or supported by monitoring studies (see chapter 3), which should be focused on the influence factor of interest and the potentially confounding factors. The studies should be thoroughly planned to ensure detection of peak concentrations and concentration profiles as well as potentially sensitive endpoints regarding physiological, ecological and statistical aspects. For this, the observation methodology and the selection of sampling sites including appropriate reference sites are crucial.

Most promising is a combination of approaches. In this way, comprehensive data packages can be generated by including high-level expertise of different approaches. For example for the future needs of marine risk assessment we want to combine our expertise in aquatic risk assessment with the know-how of marine systems and their specific ecology and ecotoxicology at the Alfred-Wegener Institute in Bremerhaven.

5 Outlook

Aquatic ecotoxicology has made enormous progresses that can support other areas of risk assessment. As exposure to chemicals and migration of organisms are definitely less complex compared to terrestrial ecotoxicology, aquatic ecotoxicology can better focus on causal effect relationships and extrapolation. For this, aquatic ecology already benefits from statistical tools developed for terrestrial objectives (e.g. TWINSPAN).

Compared to toxicological approaches, aquatic ecotoxicology has to account for far higher variability in exposure routes, physiology and ecology. At the same time, the protection aim (population versus individual) and ethical considerations (invertebrate versus vertebrate testing) enabled inclusion of much higher numbers of individuals, replicates or monitoring samples. Consequently, the statistical experience and the developed toolboxes are no longer inferior to toxicological ones, but may even provide alternatives in toxicological and epidemiological research. The advantages of exposure and non-vertebrate testing are suited to replace established vertebrate tests in toxicology, if the physiological representativity is proven, although this will be a change of paradigms. Vice versa, ecotoxicologists should include the expertise of toxicologists in physiology and risk assessment approaches at least to facilitate risk assessments for vertebrate populations.

With growing complexity and integrative needs of risk assessment, it is no longer feasible to use knowledge only in a reduced field. In the following, the starting or planned work is set into the context of risk assessment concepts that surpass the limits of (aquatic) (eco-) toxicology.

With respect to future research needs, the introductory mentioned objectives of ecotoxicology are recollected:

- The technical objective of developing, improving and standardizing methodologies, i.e., sensitive non-standard species tests, species-sensitivity distributions (SSDs), life cycle tests, modeling, specific microcosm studies, monitoring approaches.

This objective is always topical as it applies to every new conceptual and methodological development. In the future, evaluation and extrapolation techniques will gain increased importance and feed back to experimental method development (see further objectives).

- The objective of investigating causal relationships: Effects and effect compensation on different levels of biological organization, i.e. biomarkers, cells, individuals, populations or communities

The primary interactions of organisms with chemical substances are the uptake processes. For approaching causal effect relationships, increased efforts are necessary to understand the physiological processes of bioconcentration and biomagnification depending on the physico-chemical properties of the substance. These differ hugely whether the substance is dissolved, complexed by organic or inorganic matter, emulsified, bound to particulate matter or aggregated in nanomaterials. For metals, different load status can occur. The status of the substance and thus its bioavailability varies depending on chemical and physical water quality (temperature, pH, ion strength, salt concentrations, concentrations of DOM, and even turbulence). It may differ within the parameter ranges of environmental conditions (see chapter 2.4.5; see outlook 5.1).

The realization of the ecotoxicological dream of understanding causal effect relationships from molecular interactions to the population level is challenged by the high complexity of interactions and the limitation of skills to comprehensively generalize and model variability without loosing significant information. Zebrafish can be a platform to approach this extrapolation (see chapter 2.3; see outlook 5.2). With increase of understanding causal relationships, it will be possible to approach and understand microevolution processes (see outlook 5.3).

- The objective of representativity and extrapolation of research results
  - Taxonomic or physiological representation
  - Representation of communities subjected to specific routes of exposure, i.e., by municipal waste water effluents, pesticide spray drift or leaching to the groundwater

As ecotoxicological processes are physiological ones, the tests organisms and situations have to be representative of the physiology of the organism at risk. In this respect, taxonomy is only a surrogate for the unknown physiology, assuming that evolutionary close relationship also results in similar physiology. However, often adaptation to ecological contexts dominates physiological traits rather than taxonomic relationship and should be regarded important in future risk assessment (see 2.2.3 including further work). Thus, physiological representation is also important for extrapolation within a taxonomic group, specifically when using aquatic vertebrate tests for assessing the risks for human health (see outlook 5.2).

To date, routes of exposure are considered separately, as well as communities at risk (algae, pelagic invertebrates, fish or sediment dwelling invertebrates). In the future, all routes of exposure and communities may be integrated in assessment concepts, which are geo-referenced (see outlook 5.4).

- The regulatory objective: Substance evaluation or water quality evaluation
  - If substance evaluation: Context of competent legislation, i.e., for general chemicals, pharmaceuticals or pesticides
  - If water quality evaluation: Type (and use) of water body, i.e., groundwater, lentic or lotic surface waters, or marine water

The opposed directions of regulatory views can be combined when performing substance-specific regulatory risk assessments for differentiated landscapes and combining them with monitoring surveys of other legislations (see outlook 5.4).

5.1 Bioaccumulation

5.1.1 Introduction

Bioaccumulation in aquatic systems comprises of bioconcentration and biomagnification of bioavailable material. The actual OECD guideline for fish bioconcentration tests (OECD TG 305) does not systematically account for biomagnification processes, but does not exclude uptake of non-dissolved material that is complexed by DOM, aggregated as nanomaterial or adsorbed to food particles. Especially poorly soluble metals and nanoparticles need a more specific investigation of the uptake pathways of bioavailable forms (see 5.1.2).

Beside environmental health assessments, bioaccumulation in fish is important for human health and assessments of food and feed quality. These aspects have been addressed by tissue-specific analyses in OECD 305 tests, which scientifically is not the intention of a bioconcentration study. The revision of the directive 91/414 includes fish feeding studies to perform metabolism research and residue analysis. Ongoing discussions try to separate or combine objectives and methods, where appropriate (see 5.1.3).

OECD TG 305 is not only criticized for not properly addressing biomagnification and for poorly meeting needs of residue and metabolism studies, but also for being an animal test consuming too many individuals. Proposals of reduction of animals, refinement of test protocols (see 5.1.3) and replacement by different approaches are discussed (see 5.1.4).

5.1.2 Bioconcentration versus biomagnification for poorly soluble metals

Metal bioconcentration tests tend to result in bioconcentration factors (concentration of a chemical in fish divided by its concentration in water = BCF) that depend on the exposure concentration (ref.). This is not in line with bioconcentration theory, which assumes uptake and elimination being dependent on the concentration in water and consequently the BCF being independent of the exposure concentration. Whereas organic chemicals sometimes increasingly accumulate with increasing concentration, if metabolic potentials are limited, for metals often a higher BCF is determined at lower concentrations. The deviating results of metal bioconcentration tests may be a consequence of metal regulation processes, which actively enhance the internal concentration of essential metals at low ambient water concentrations and reduce it at high ambient concentrations. Other metals may be co-regulated.

Results of **bioconcentration tests with two poorly soluble metal compounds** pointed to an alternative interpretation. There is evidence that non-dissolved material was responsible for the findings. Independent GLP studies were performed with two different metals of the same element group according to OECD TG 305. The tested metal compounds are poorly water-soluble. In pre-tests the conditions to achieve the highest dissolved metal concentrations possible (operationally defined as metal concentrations in 0.45 µm membrane filtered water samples) were identified. The stock solutions were prepared under acidic conditions and dosed to the purified drinking water, resulting in a pH in the test chambers stocked with fish of around seven. There was no visible precipitation or dispersion. At the end of the uptake phase of the second test, surplus fish were dissected in the intestines, the outer parts (skin, gills and head) and the inner parts to investigate the particular accumulation in these tissues. Furthermore, a feed adsorption test investigated the metal accumulation in feed. Both tests resulted in characteristic dependencies of the BCF (Figure 52).

When separating the inner tissues, the intestinal tract and the outer parts, it became obvious that there was no accumulation due to adsorption to outer parts, neither due to bioconcentration via the gills in the inner fish (Figure 53). The only parts showing bioaccumulation dependent on the exposure concentration were the intestines. From the guts, which are a turned in inverted external part of the fish, there is obviously no significant uptake into the inner fish. The feed adsorption demonstrated a considerable accumulation in the feed. However, the results could only explain 10% of the metal concentration in the guts at the low treatment and 2% at the high treatment.

For the two investigated poorly soluble metals, the adsorption to cell membranes and mucus via ingestion seems to be the only significant uptake pathway, being most relevant for particles. This can either be metal adsorbed to food particles or metal nanoparticles ingested with the water, as the 0.45 µm filtration is not able to differentiate between dissolved and nanoparticle fractions. The “BCF” for the total fish (Figure 1), which in fact is not a bioconcentration factor, cannot be explained by the presence of feed residues, but with concentration dependent association of the metal with the intestinal surface layers, which can be interpreted as a hint to the ingestion of nanoparticles. The material associated with the guts does not seem to be bioavailable.

A more thorough differentiation of the uptake of dissolved metal and metal nanomaterials is going to be analyzed in different public projects with silver nanoparticles. A fish early life stage test was performed in a test system optimized for a homogeneous exposure to nanoparticles: Aquaria with sufficient volume and agitation of the water body were equipped with stainless steel cages serving as pseudo-replicates for statistical evaluation. The test suspension was renewed weekly. Beside early life stage toxicity, accumulation of silver are going to be investigated in juvenile fish, The fish will be dissected in 1) skin/gills, 2) guts, 3) filets and 4) remaining tissues, each part analyzed for total silver concentration.

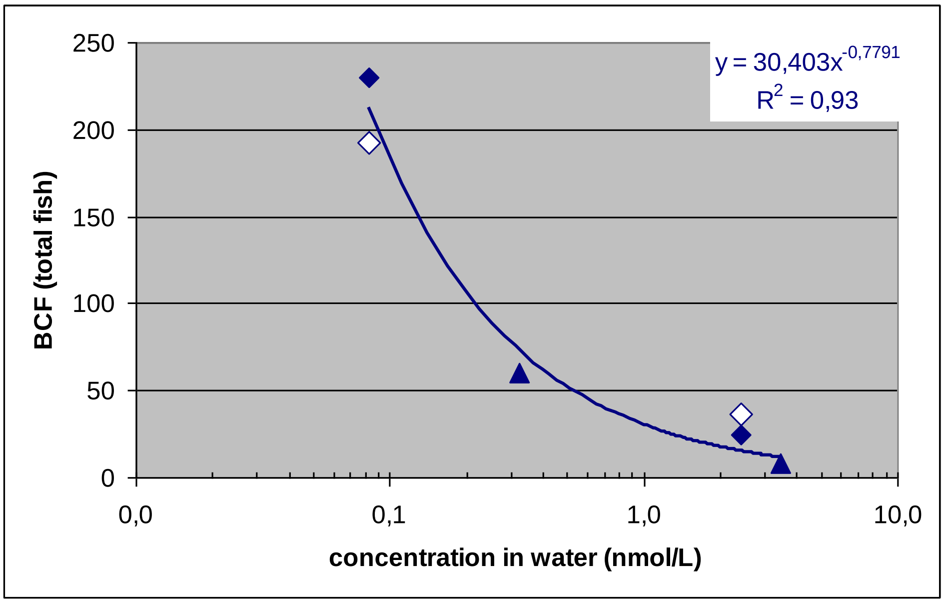


Figure 52: Concentration dependency of the BCF. Data of independent tests with two metal compounds (1: diamonds, 2: triangles) of the same element group. Open diamonds: Total fish data of dissected (additional) fish

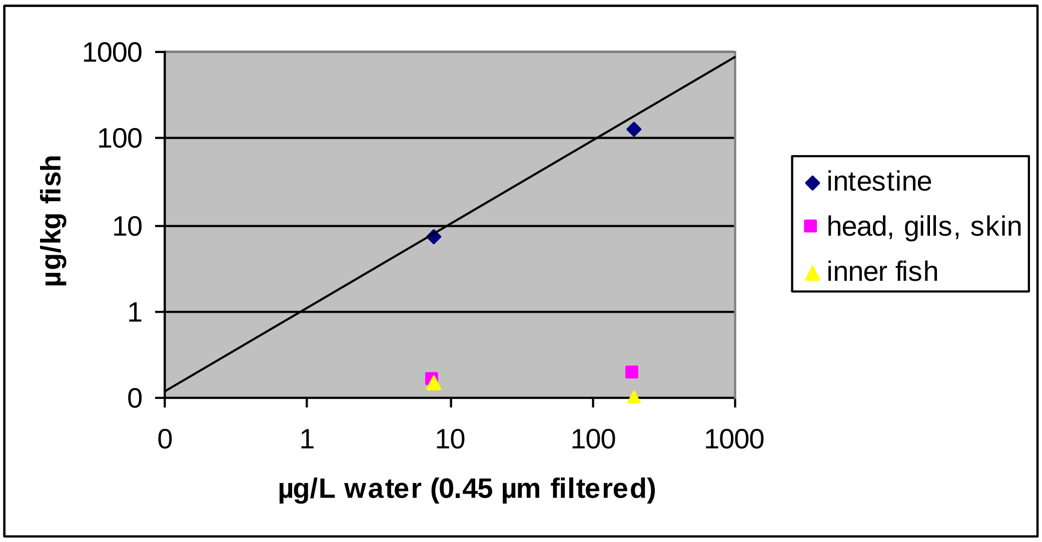


Figure 53: Accumulation of a poorly soluble metal in different fish tissues

5.1.3 Revision of OECD TG 305

Already in 1996, industry presented a comprehensive proposal for the revision of OECD TG 305. The revision is now part of the actual work program, with Germany, UK and The Netherlands being the lead countries. In 2008, UBA as national coordinator organized a discussion meeting with an invited group of German experts to discuss needs of and proposals for the revision.

Test objectives. The primary clarification should be on the objectives making use of the study results. Environmental risk assessment of substances needs the bioconcentration factor for the whole fish, the clearance time including a potential residue plateau and sometimes information on the identity of main metabolites. Consumer risk assessment needs information of residues in edible parts of the fish, which may depend on cultural particularities. For this, separate concentration factors are determined for edible parts (e.g. filet) and viscera; metabolites are quantified separately, too (e.g. US-EPA guideline 165-4: identification of metabolites >50 ppm in the edible parts).

Number of test concentrations. When meeting bioconcentration theory, the BCF is not dependent on the test concentration. Thus, one test concentration without toxic effects is sufficient for the determination of the BCF. The US-EPA guidelines 165-4 and 72-6 are based on testing one concentration. OECD TG 305 prescribes the use of two test concentrations differing by a factor of ten to prove validity of the test. If the two concentrations yielded different results concerning kinetic parameters and the BCF, either at least one concentration was not set up correctly (inappropriate exposure, toxicity), or the bioaccumulation of the test substance does not follow bioconcentration theory (concentration in fish is proportional to concentration in water).

- The test substance can be metabolized up to a limited dose. When the metabolic potential is exploited, the BCF increases with increasing concentration in the media (IME GLP study example). For a profound environmental risk assessment, the knowledge of and relation to the range of environmental relevant concentration is necessary.
  If there is suspicion of metabolic limitations in a sublethal concentration range, a bioconcentration study should be performed beyond OECD TG 305, based on an appropriate number of test concentrations.
- The test substance is regulated or co-regulated as it is evident for natural elements (IME GLP study examples). In this case, the concentration of the test substance in ish tends to be constant, irrespective of the concentration in the media, if it does not exceed the norm of reaction. Consequently, the BCF increases with decreasing concentration in the media.
  If there is evidence for a regulation or co-regulation of the test substance, as it clearly is for essential elements, a bioconcentration study is inappropriate.

Beside ethical considerations (high numbers of vertebrates are sacrificed), there are often analytical problems associated with the necessity of two test concentrations differing by a factor of ten. They mainly occur with substances of high toxicity and/or poor water solubility. As testing of partly insoluble material results in incorrect BCFs, water solubility in the test media used should be known before the test and can more easily met by testing one than two concentrations.

Lipid content. The BCF is based on fish wet weights. As it depends on the lipid content of the test fish, the lipid content has to be measured and reported. Sometimes the BCF also relate to the lipid content. As the regulatory community is not familiar with these results, industry proposed normalization to a lipid content of e.g. 6%. The IME (Christian Schlechtriem) proposed a methodology which will be part of the revised guideline.

Fish metabolism is not explicitly included in usual bioconcentration models. Thus, the definition of the plateau (steady state) does not account for the most usually observed decline of the bioconcentration curve. As the BCF should be protective for all fish species including weak metabolizers, the plateau should be defined as state without further increase of bioconcentration. The actual version of the guideline recommends to identify metabolites in “chemicals such as pesticides”. This should be extended to further classes of substances, especially pharmaceuticals, as the main metabolite often is the active substance. However, the general assumption is that metabolites are more hydrophilic and less toxic compared to the parent compound.

Metal bioconcentration. As uptake and distribution mechanisms as well as homeostatic regulation is different for metals compared with organic chemicals, metal accumulation is not covered by bioconcentration theory (see 5.1.2). Thus, OECD TG 305 is not suited for testing metals, whereas valid results can be achieved for organometal compounds.

Feeding and test duration. At a study duration of 28 d fish have to be fed. This causes problems with bioconcentration theory, as adsorptive test substances will adsorb to feed particles and be taken up via feed rather than via the water phase. A reduction of the study duration to one week enables the conduction without feeding. For most of the test substances, 90% of the plateau is reached after one week. The BCF should be calculated as kinetic BCF. A reduction of the test duration also is helpful when testing substances with difficult physico-chemical properties, needing high efforts to manage the test conditions. (e.g. high volatility; stable only in a small range of pH as for some pharmaceuticals).

Fish Feeding protocol. If the test item properties reduce bioavailability in the water phase, but induce biomagnification via the food chain (e.g. log $P_{OW} > 6$), bioaccumulation via bioconcentration is less relevant than via feeding and should be addressed by specific feeding studies. The actual fish feeding study protocol that will be included in the revised OECD TG 305 is going to be validated. IME participates in the validation, evaluation and discussion of regulatory consequences.

To achieve a scientifically more correct assessment of risks for bioaccumulation, the uptake routes via bioconcentration and dietary exposure should be investigated independently. Consequently, the not addressed uptake routes should be excluded, resulting in non-feeding protocols in a revised OECD TG 305. A clear definition of the bioconcentration factor (BCF), the biomagnification factor (BMF, per trophic level?) and the bioaccumulation factor (BAF) as an integrative measure (referring to the concentration in water or sediment?) will be necessary. The real situation will be a combination, but can only be assessed for the different fish species by distinction of the processes and addressing their individual variances. To date there is no experience on comparisons between results of fish bioconcentration test (BCF) and dietary exposure (BMF). The contribution of the different exposure routes may vary enormously in different fish species, depending on species-specific diets (e.g. algae/macrophytes, zooplankton, benthic organisms, fish) habitats (e.g. sediment, lentic waters, lotic waters) and oxygen needs (see 2.2.3). Metabolic capability should also be considered, as it is an important driver for bioaccumulation (Preuss 2006). To facilitate risk assessment schemes, it is advisable to select scenarios representative for fish species especially susceptible when regarding classes of physico-chemical parameter values. This may i.e. be rainbow trout for water soluble substances (high bioconcentration due to high water exchange at gills, weak metabolizer due to low water temperature and salmonid physiology) and eel for poorly water soluble substances, e.g., of high $\log P_{OW}$, (high biomagnification due to feeding on sediment organisms and a high lipid content).

5.1.4 Fish metabolism and PBPK modeling

Dietary fish metabolism study. Due to the growing importance of crops in the diets used for aquaculture, the new EU guideline 1107/2009/EC for the registration of pesticides includes the requirement for residue and metabolism studies in fish. IME (Christian Schlechtriem) developed a method for metabolism studies, which was explored with carp and rainbow trout. As metabolism studies should be dietary studies only (also for substances with a log $P_{OW} < 6$) and the fish should be of sufficient size to gain material for metabolite identification, the study should be sepatated from OECD TG 305 studies.

PBPK modelling. Beside the uptake routes, the distribution of active substances in fish should be considered in a more realistic and thus complex way as done before. Making use of the knowledge of (human) pharmacology and toxicology enables the development of physiologically based pharmacokinetic models (PBPK) for fish (Abbas and Hayton 1997), In this respect, again it is advisable to select few physiological extremes to account for the variability of different fish species.

Fish metabolism can considerably vary between species (see chapter 2.3.6.3). The metabolic capacity depends on temperature and metabolic capability, the capability on the (iso-) enzymatic equipment. Residue and metabolism studies needed for the assessment of consumer risk require ^{14}C-labeled test substance and the inclusion of species being consumer-relevant as well as sensitive concerning the bioaccumulation potential. In this respect, beside rainbow trout (see chapter 5.1.3) a marine flatfish with high lipid content seems appropriate.

Regarding animal welfare issues, bioaccumulation tests on whole vertebrate organisms should be refined (see chapter 5.1.3), reduced or replaced. For the latter, the important role of metabolism is focused on when developing non-animal test for bioaccumulation. The most promising results so far are yielded from rainbow trout primary hepatocyte tests (SETAC Bioaccumulation Advisory Group 2008). When combined with PBPK modelling approaches, it should be possible to gain information sufficient to replace regulative bioconcentration tests. At the same time, inclusion of hepatocyte tests and PBPK information from different fish species enable a more precise estimation of variability than to date. However, a lot of applied basic research and validation work will be necessary to finally replace regulatory OECD TG 305 tests. As Fraunhofer IME is experienced in cell culturing, bioaccumulation testing (OECD TG 305 tests and field simulation experiments) fish metabolism and modelling, we will take part in the development. We started a PhD work on fish metabolism studies including primary hepatocytes in cooperation with Helmut Segner, University of Bern, Switzerland.

5.2 From physiology to environmental effects

5.2.1 Introduction

Understanding causal effect relationships from molecular interactions to the population level has to cope with the high complexity of interactions on and between the levels of biological organization. This is only possible by comprehensive generalization. However, the methods of generalization should always be transparent and appropriate for its objective, i.e. the reason for and ways of excluding information should be justified. At the same time, the variability of potential causalities should be clear at each level. Consequently, the possibility of extrapolation representing the real world is very limited due to limits of knowledge and due to the stochasticity of real interactions, whereas it is highly probable to match a result looking like a part of the distribution of real world situations by chance.

With these restrictions, zebrafish seems to offer the most promising approach for an extrapolation from molecular effects to effects on populations. Its genome is characterized and the database on toxic effects is growing. We intend to develop a non-target screening tool for MoA-specific responses in zebrafish embryos and to investigate molecular effects. At the same time, we try to link these effects to morphological changes in tissues and organs (see 5.2.2).

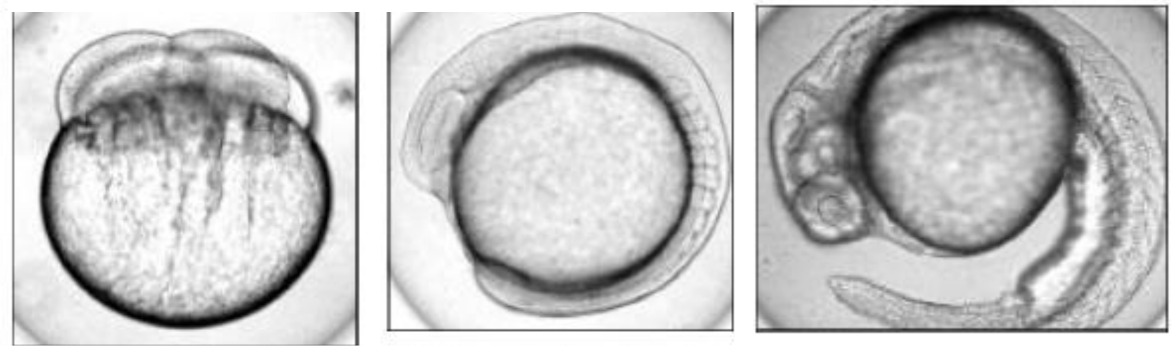
On the individual level, we are able to run every state-of-the art method in fish toxicology up to full life cycle and multi-generation tests. For endocrine disrupting chemicals, we have a sound overview of population-relevant effects on individual performance (see chapter 2.3). The understanding of effects can be improved by including modeling of uptake, distribution and metabolism (see 5.1). Different kinds of population models may be combined for the extrapolation to population-level effects (see 5.2.3).

5.2.2 UNIFISH – a universal experimental tool for effect identification in fish

In January 2008, we successfully applied for a Fraunhofer Attract-funding of a research group of two scientists, three PhD students and a technician for 5 years.

The scope of the group is the development of a universally usable screening tool for toxicological effects that ideally should be able to detect all potential effects on vertebrates. For this, zebrafish embryos seem to offer a wide range of advantages towards other screening tests based on molecular or cellular tests, as they enable:

- integration of a total organism with all basic metabolic functions,
- detection of potential effects especially sensitive during development, as it passes different levels of biological organization (cell, tissue/organ, organism), each level providing potential targets of effect manifestation,
- performance as high-throughput test (small scale in 96-well plates, short-term in 48 h),
- the linkage between high-resolution 3D-image analysis and a molecular effect profile.



1h 12 h 24 h

Zebrafish embryos can serve different protection goals (see also Nagel 2002).

1. Ecosystem health. As representative organism for aquatic secondary consumers, zebrafish is an accepted test species for ecotoxicological hazard assessment, using acute test (OECD TG 203), embryo/sac fry tests (OECD TG 212), early life stage tests (OECD TG 210) or life cycle (partial or full)/two-generation tests (see 2.3).
   As unhatched eggs legally are no vertebrates capable of living, testing effects on fish eggs is not under the responsibility of the animal welfare act and categorized as non-animal testing. As zebrafish embryos are accepted to be comparably sensitive to effluents than the former test organism, the golden orfe (*Leuciscus idus*), the zebrafish egg test (DIN 38415-6, 2001) is one of the first alternative test systems implemented in a national regulation (German waste water dues law, Bundesgesetzblatt 2005).
   To date, there is an initiative led by Germany to validate and adopt the fish embryo toxicity test as an OECD test guideline (OECD 2006b). The potential replacement of the acute fish test for hazard assessments by this guideline needs sound investigation of the limits of uptake through the egg membrane as well as of the developmental stage of structures susceptive for toxic action.

2. Human health. When regarding the requirements resulting from the new EU chemicals registration policy 'REACH' (Registration, Evaluation and Authorization of Chemicals), there is need for an alternative screening method prior to the conduction of teratogenic test according to OECD TG 414. The applicability of zebrafish embryos was already checked (Bachmann 2002), but needs further validation.
   The zebrafish has become one of the most popular model organisms in basic research, toxicity testing and drug discovery. Because vertebrates are genetically very similar, the zebrafish is a very promising research tool for studying human disease (Dooley and Zon, 2000). This is reflected by recently funded European Union projects such as ZF-MODELS (Zebrafish as a Model to Study Human Development and Disease; www.zf-models.org). The potential of the zebrafish in drug development and as a basic research tool has been recognized by many researchers, and a number of start-up companies (Zygogen,USA; ZF Biolabs, Spain; DanioLabs, UK, acquired by VASTOX; ZF-Screens, Netherlands; Biotecont, Hungary; Phylonix, USA) have been established to work exclusively with this model. Compared to the activities of these companies, UNIFISH will focus on a combination of automated morpho-metric and molecular endpoint analysis.

The screening method to be developed by the started work is based on the principles shown in Figure 54. The main methodological challenges for the image analysis part are the virtual orientation of recorded images for comparison purposes and the identification and parametrization of structural responses. For the toxicogenomics part, the main challenge is the appropriate selection of genes for the microarray according to the principles mentioned in 5.2.1 (justified reasons for exclusion under the objective of non-target screening). We intend to analyze the response patterns statistically by using multivariate methods used for ecological site classification (see 3.4). Thus, MoAs will be classified according to the most typical response patterns, based on correspondence analysis. At the same time, the indicative power of single responses will be ranked.

The central work will be the generation of a comprehensive reference data base of response patterns of as many different MoAs as possible.

The potential applications for the non-target screening tool are screening for potentially adverse effects as well as for potential new active substances. In a tiered approach of toxicological hazard assessment (e.g. for the necessity of conducting teratogenic assays on rodents), it may be sufficient to focus on the image analysis approach. A high-throughput screening during active substance development on adverse effects should include both. When screening for potential active substances, the focus is on the toxicogenomics approach. The approach is especially valuable for the investigation of natural substances, e.g. in extracts.

Exposure of fish embryos

Image analysis	Toxicogenomics
Close recording (e.g. hourly) Automatic comparison of parameters with control	after 48h homogenized, RNA extraction Microarray, substration of control response
Distinction pattern = response pattern	Remaining responses = response pattern
Automatic comparison with database on response patterns (known MoAs)	Automatic comparison with database on response patterns (known MoAs)
Result: n% probability of MoA x; m% probability of MoA y…	Result: n% probability of MoA x; m% probability of MoA y …

Comparison and interpretation of results

Figure 54: Principles of the UNIFISH screening tool

UNIFISH may be most valuable for testing matrices or media containing different and especially unknown substances, such as environmental samples, products, or food and feed samples. It may find hints to specific MoAs induced, and thus enable hypotheses on the presence of significant amounts of adverse substances. This effect-related analysis of samples can guide specific chemical analyses. A prerequisite for this kind of application are methodologies of enrichment of water samples and extraction from soil or products that are as efficient as possible (no losses of formation) and compatible with fish embryo testing.

Besides the screening purpose for (eco-)toxicological objectives of regulation or survey, the results can be used for a number of basic research issues. The research on indicative molecular biomarkers will be supported by a sound database and tools for a statistical evaluation of predictive potential. The statistical classification of MoAs by response patterns and ranked indication potential of single responses is a useful tool for supporting bioinformatic approaches.

The PhD and post-doc work planned in the program is organized in classes of effects or substance properties, which are related to specific regulative contexts and extrapolation needs. For all MoAs included in a work package, several representatives inducing weak to strong responses are to be tested. The methodologies will have to be optimized for a detection of characteristic response patterns (parametrization of structural effects, selection of microarray genes).

1. Endocrine disrupters: All known MoAs (for sexual endocrine effects see chapter 2.3.6); inclusion of medaka (extrapolation potential between fish species); contribution to osteoporosis research (extrapolation potential to mammals for the evolutionary conservative steroids)
2. Pharmaceuticals: Selection of relevant MoAs, comparison to responses in mammals (general extrapolation potential from fish to mammals)

3. Pesticides: Selection of MoAs relevant for fish other than endocrine disruptors, comparison with insect toxicogenomics (Fraunhofer IME group planned at the University of Gießen: extrapolation potential vertebrates : arthropods)
4. Metals and nanomaterials (and further substances with potential barriers for uptake through the cell membrane, e.g. by molecule size or superlipophilic properties) (extrapolation potential to cute fish tests)
5. The work may be added by similar investigations on fish pathogens (exposure by microinjection).

I wish to thank Stefan Scholz, UFZ Leipzig, for contributing his ideas and experience in the field of toxicogenomics, and for writing the main parts of the application proposal. After his decision to remain at the UFZ as head of department, Martina Fenske took over responsibility for the finalization of the proposal and the UNIFISH working group.

5.2.3 Predictive modeling / risk assessment

When a molecular effect is identified by a specific response pattern in fish embryos, the relevance for effects on individual life performances is to be shown. This is straightforward when observing morphological or even lethal effects on the fish embryo. The extrapolation to individual juvenile fish of the same or different species can be based on empirical testing. However, using the existing data and toxicological experience for generating models of uptake, distribution and metabolism (see 5.1.4) helps to reduce future vertebrate testing, if the extrapolation is possible via modeling of differences in uptake and internal exposure. There is still need for hypotheses on the relation between molecular effects and individual life performances for many MoAs. For exemplary work, it seems easiest to focus on sexual endocrine disruption, as there is good information on molecular responses and individual effects from endocrine-specific endpoints and assays (see chapter 2.3).

In September 2008, global fish ecology and fish ecotoxicology experts as well as statisticians very invited to a workshop on fish full life cycle testing for the hazard assessment of endocrine disrupting chemicals in Palma de Mallorca, Spain. After agreement on important endpoints to be regarded in fish full life cycle tests, they were ranked for significance by means of a multi-criteria decision analysis (MCDA). The criteria included costs and practicability criteria, ethics, statistical power and the risk of false negatives or false positive responses, relevance for populations and the environmental situations, regulatory relevance, sensitivity and public views (ref. 26). The discussions between fish ecologists and ecotoxicologists pointed to the importance of evaluating integrative population relevant endpoints (total early life stage survival, sex ratio, growth) rather than sensitive ones, which might be more indicative.

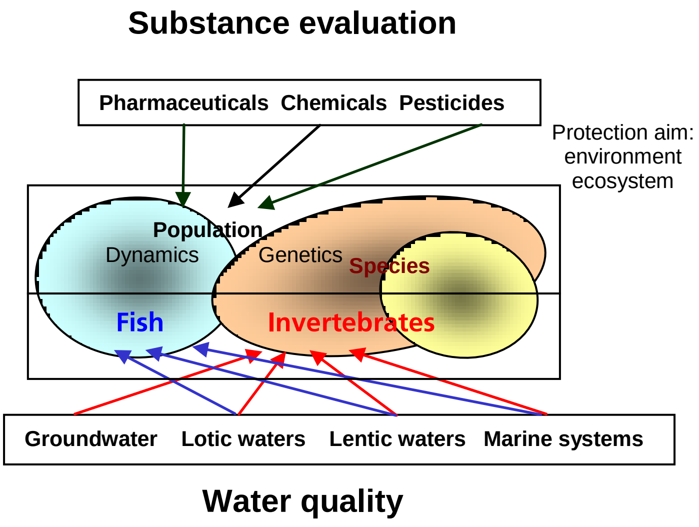
With respect to modeling, individual based models (e.g. on zebrafish and guppy, ref. 2),7)-9)) are able to include effect data generated in life cycle tests. However, they are very specific for the species and situation modeled and the extrapolation capability is not explored. They should be compared with more general stage or recruitment models, which are available for many fish populations but which are not capable to translate specific effects of chemicals. Thus, there seems to be necessity for a link between the different types of models, the individual models for scientific understanding and the stage or recruitment model results for regulative decisions. To date, Udo Hommen (IME) directs a Marie Curie PhD work performed by Lara Ibrahim on fish population models.

5.3 Population genetics and landscape level risk assessment

5.3.1 Introduction

As pointed out in the previous chapters, substance evaluation and media quality evaluation are in the responsibility of different legislations. With respect to ecotoxicology, they interlink through the derivation and control of quality objectives.

When looking at the common good of protection, the aquatic community, the maximum achievable aim of ecotoxicological substance-evaluation concepts is to predict risks for the dynamics of populations, maybe meta-populations in the future when regarding dispersal and colonization patterns dependent on habitat connectivity in the landscape.



These predictions can be compared with monitoring results and used for interpretation of community structures (Figure 55). For the analysis of impacts on the aquatic community, an ecological understanding of community structure is a prerequisite, including knowledge of the determining habitat-specific environmental factors, which may confound predicted effects of chemicals.

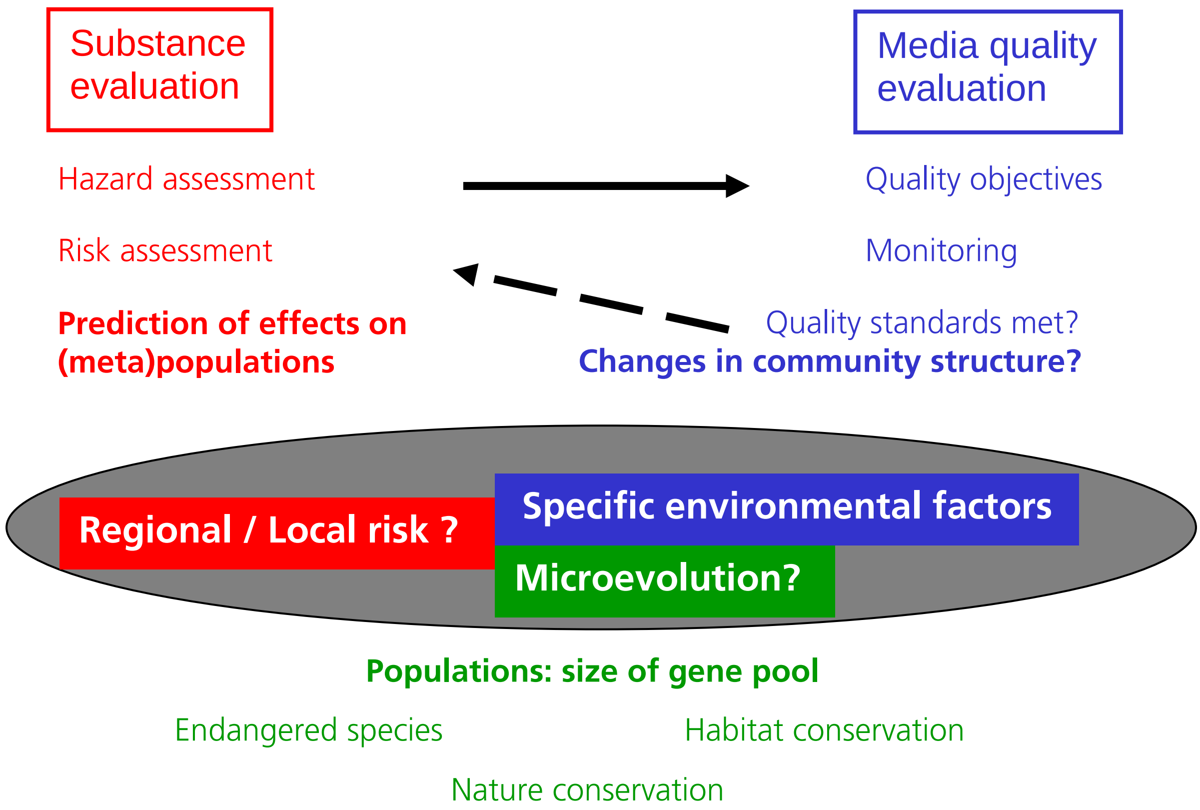


Figure 55: The population as central unit of protection concepts

The objectives, concerns and methods of nature conservation approaches are different. The main objective is defined more clearly, as it is biodiversity. In a first instance, this is species richness, which is dependent on the number of ecological niches, guaranteed by a high diversity of habitats. A local community consists of populations as local subunits of species, characterized by genetic exchange. For the maintenance of the potential to adapt to environmental challenges, the gene pool of a population should be wide, containing high variability in alleles, resulting in high degrees of heterozygoty and low influence of genetic drift. This can only be achieved by a high connectivity of the population-specific habitats. Thus, the main objective of nature conservation is habitat conservation, especially focusing on habitats for endangered species. The scientific aim of nature conservation is the understanding of specific patterns of microevolution, which can be approached by investigations of population genetics (Figure 55). In that view, chemical stressors are minor influence factors beyond others. However, for understanding chemical effects in a context of landscape ecology, knowledge on population dynamics and population genetics should be linked.

5.3.2 Population ecotoxicology – a field of transition between assessing the risk of pesticides and nature conservation

The population is the integrative biological level where predictive and retrospective objectives and methodologies of ecotoxicology and nature conservation meet. It is also the level at which processes of (micro-)evolution become manifest. The main factor influencing population genetics is intraspecific competition. The consequences of directed selection by anthropogenic stressors and gene flow between impacted and not impacted populations have to be taken into account when assessing the risk of chemicals applied to the environment. This requires regional and even local approaches.

5.3.3 Risk assessment for pesticides and nature conservation

The central protection aims with respect to dangers resulting from the application of plant protection products were mentioned in article 4 of EU directive 91/414 and, more precisely, in § 1.4 of the German Plant Protection Act (BML,1998): They are human and animal health, and the "balance of nature". The "balance of nature" is defined in § 2 as its compartments soil, water and air, the animal and plant species and the interactions between all. As the species are explicitely mentioned, structural aspects without relation to function dominate the ecotoxicological risk assessment discussions, at least when animals are in the focus (ref. 31).

Risk assessment within the scope of the notification of plant protection products is mainly based on predictive approaches: effect and threshold concentrations indicating the sensitivity of test organisms to a defined test substance (hazards) are compared with predicted concentrations of a potential exposure in the field. The toxicity data are gathered from single species tests on individuals (e.g. fish, invertebrates, plants) or growing populations (e.g. bacteria, algae) in lower tier tests up to age structured test populations of different species in higher tier community level tests (Campbell et al. 1999, (ref. 12). There is increasing effort to model species that cannot be represented by age-structured test populations and tested for more than one generation, like vertebrates. By modeling the population dynamics, available relevant information about the population ecology of a species can be introduced in assessing the age and performance specific effects on populations from results of testing individual organisms (e.g. ref. 2), 8), Dülmer 1998). The population level is the ultimate level of integration that might be assessed concerning realistic and relevant effects predicted from suited ecotoxicological tests (Rudolph and Boje 1986). The limitations of predictive population ecotoxicology, e.g. by the lack of knowledge of kind and uncertainty of influence parameters, are clearly seen but are not subject of this outline.

Species protection as a part of nature conservation is regulated by different legislation. A central tool is the Red List of endangered animal and plant species (BfN 1998). All species are classified for different degrees of threatening. The main classification criteria are

- actual abundance (number of populations; mean population size; distribution of populations in space and time)
- historical development (in state of increase or decrease, dynamic of change; dynamic of habitat loss/win; loss/win of site or habitat-specific populations)
- prediction of development in future (realistic worst case of continuing development without considering constantly necessary conservation efforts; threatening by direct (local) and indirect (regional, global) anthropogenic influences
- biological risk factors (low reproduction; short survival of adverse conditions; low migration and colonization capacity; adaptation to very specific conditions; close obligation to another declining species; danger of hybridization with a closely related and much more abundant species).

Data acquisition within the scope of this classification is mainly retrospective by detecting and assessing a state in the field: The abundance of (especially rare) species is observed. It is attempted to link deterioration with possible causes. In this approach, the lowest investigated level of integration is the population level: Extinction manifests itself firstly by the extinction of local populations. Identifying genetic diversity of an individual population of an endangered species is used to indicate genetic exchange with neighboring populations, the extent of isolation and – consequently – the potential risk for the population (Reh 1991; Reh and Seitz 1990; Seitz and Loeschke 1990).

The approaches of regulatory ecotoxicology and nature conservation are entirely different with respect to focus, methodology, consequences for legal execution and involved expert groups, partly opposing or at least ignoring each other. However, they overlap concerning the protection aim and concerning the population as operationally final endpoint.

The importance of intraspecific competition in case differing genotypes perform different sensitivities towards plant protection products

Intraspecific competition causes the strongest pressure on ecological maintenance, because individuals of the same species occupy the same ecological niche and compete concerning all resources. This situation, together with the multitude of other factors of influence, leads to a more or less steady state of genetic diversity being specific for each population. It can be described by the number and frequency of allele genes or by the extent of heterozygoty. A state of high genetic diversity enlarges the norm of reaction of a population on natural and anthropogenic stressors and thus the potential of survival (Calow and Berry 1989; Beaty et al. 1998). A shift in the intraspecific competition triggered by a distinct influencing factor - for example yearly applied pesticides / active ingredients with the same mode of action, which affect the performances of some genotypes more than those of others – causes a shift of the allele frequencies within a population potentially resulting in the extinction of some alleles. Thus, the genetic diversity of a population may deteriorate.

Besides that directed pressure of selection, some further processes are important for microevolution within short time intervals: gene flow via migration and genetic drift (stochastic extinction of alleles of low frequency) (Wilson and Bossert 1973). Migration results in the introduction of genotypes of a population into a neighboring population. On one hand, locally disadvantaged alleles may be steadily reintroduced this way, diminishing the consequences of directed pressure of selection in mobile species. The cultivated landscape in central Europe, however, provides only few refuges being large enough for maintaining sensitive genotypes. Thus, on the other hand, the opposite scenario may apply: Especially mobile species also tend to introduce the less sensitive and more deteriorated genotypes of sublethally impacted populations into untouched populations. If the relation between cultivated and pristine areas is disadvantageous, the more sensitive genotypes become relatively rarer even in the pristine areas. The allele frequency might decrease below a threshold level, where the process of genetic drift resulting in extinction becomes more probable. Subsequently, the genetic diversity of species inhabiting the cultivated landscape might decrease.

Beaty et al. (1998) investigated the diversity in the 16S rDNA of mayfly *Baetis tricaudatus* populations. The objective of the study was to retrospectively identify and interpret population genetic responses to site specific heavy metal contamination in a river. Genetic diversity inversely correlated with metal exposure. A minimally exposed tributary introducing softer water in the contaminated river seemed to cause further genetic deterioration: The effect of increasing bioavailability of the metal load of the main stream superimposed the dilution effect. At sites far downstream the spot contamination, genetic diversity increased again, correlating with the recolonization probability, driven by the influence of uncontaminated tributaries. These findings support the hypothesis that anthropogenic chemical stressors can cause deterioration of population genetics. Whether gene flow via migration is able to compensate the loss of alleles is depending on the proportion between contaminated and uncontaminated sites.

5.3.4 Implications for future work

The inclusion of population genetics within the scope of risk assessment for pesticides is close to the scope of species protection. In the light of the consideration of regions within European legislation on plant protection products, the objectives may become more compatible. In future, even a local view might be indicated. In any case, a scientifically more detailed elaboration of respective common criteria for, and remaining differences in protection aims is required to achieve more consistencies between plant protection and nature conservation. I hope that this will lead to a clearer view on risks of chemicals for biodiversity, a more precise definition of the protection aim of substance evaluation and a localized risk assessment. The result might be a more diverse pattern of use intensity, depending on the landscape-specific risks and benefits instead of intending to protect everything and ending in disregarding specific risks as well as in wasting obvious benefits.

5.3.5 Participation in a planned German center for biodiversity research

In February 2008, the German research society DFG invited for a symposium in the science forum, Berlin, about biodiversity research in Germany. The tasks of the symposium were the inventory of German biodiversity research, the assessment of strengths and weaknesses, the listing of development needs and the induction of measures for improving structure and funding. The invited delegates were scientific representatives of the Universities, and of the German research associations Max Planck (MPG), Helmholtz (HGF), Leibnitz (WGL) and Fraunhofer (FhG). As the IME is the only Fraunhofer Institute focusing on ecology, I was delegated as representative of the FhG.

The strengths were identified to be the scientific collections and taxonomic expertise, making Germany worldwide leading in recording and describing biodiversity, particularly in tropical regions, in the Antarctic and the marine ecosystems in general (research ships Polarstern and Meteor). Weaknesses compared to activities of other research countries are terrestrial and freshwater microbiology. A threat is the trend to reduced taxonomical education at the universities. At the same time it was underlined that an appropriate network between the collections is completely missing. Thus, the use for central, hypotheses-driven research is nearly impossible. Consequently, Germany is strong in macro-evolutionary research, but there are deficits in micro-evolutionary research, e.g. in population genetics.

The primary need is the organization of a network of databases of the scientific collections and the inclusion of the university data and samples. Hypotheses-driven meta-evaluations of the centrally accessible databases should be enforced. At the same time, the population genetic biodiversity research should be encouraged to identify the drivers of biodiversity in stress ecological contexts, as this is the main political expectation.

The decision board of MPG, HGF, WGL and the university director conference was informed about the symposium results to agree on future strategies. A German forum for biodiversity was founded to guide the further development.

The most urgent action is the formation of a common German Centre for Biodiversity, which will receive scientific and financial input from MPG, HGF, WGL and FhG and will be located at an appropriate university. This centre will be responsible for the central tasks of organizing a Germany-wide database network. In fact, it will represent only the central nod of a decentralized network, as biodiversity essentially is a decentralized research topic.

As the participants in the German forum for biodiversity are mainly from basic research, the Fraunhofer IME may contribute applied research topics. These could be influence-related monitoring concepts (see chapter 3.4), geo-referenced assessments (see chapter 5.4), applied population genetic research (see this chapter), the use of the environmental specimen bank located at Schmallenberg (storage of soil samples), or the development of biosensors (department of environmental and food analysis). Politics needs such projects and we are able to contribute ecological and statistical expertise as well as useful technologies (cryobanking, biosensors).

5.4 Georeferenced risk characterization and prediction

5.4.1 Introduction

In Germany in the 1970s and 1980s, the environmental risk assessment of chemicals was dominated by environmental chemistry. Even the terminus ecotoxicology was defined as a chemical one (Parlar and Angerhöfer 1995); ecotoxicology was regarded as substance-specific property, measured in standardized, mostly acute, ecotoxicity tests. During the last two decades, the risk assessment procedures becoming more and more internationalized, the physiological and ecological properties of responding organisms and population-relevant effects were in the focus of ecotoxicology, associated with questions of representativity of test organisms, recovery and extrapolation to real communities.

Now it is evident that these crucial aspects of risk assessment can no longer be represented by worst-case scenarios, the derived risk being protective for all communities in any landscape of the regulated entity of states. Exposure and potentially affected populations are geographically distributed, differing in a way that worst-case scenarios would lead to unacceptable risk in most cases. In the EU pesticide regulation, the regional exposure is accounted for e.g. by the FOCUS scenarios, which, however, only are able to state safe use of a pesticide application in a represented geographical sub-region without considering geographical variation of effects. Geographic information systems (GIS) are widely available now and will be used for environmental risk assessment of chemicals in the near future. We are part of a consortium developing a GIS-based tool for pesticide risk assessment (see chapter 5.4.2). By applying GIS it will be possible to locally include land use patterns, topographical structures and microclimatic aspects, revealing confounding influence factors like riparian vegetation, buffer zones, proximity to e.g. roads or waste water treatment plants, and causing local distribution of potentially responding populations. When GIS-based risk assessment procedures are available, the recovery potential of a regional population can be assessed by comparing the local distribution of e.g. pesticide use patterns with that of local sub-populations (see chapter 5.3.3). This opens the vision for a harmonized characterization and prediction of risk for all influences on communities, no matter whether they are now regulated by REACH, the pesticide regulation, waste or waste water regulations, or subject of the habitat conservation legislation and nature conservation. The full substance-specific information about chemical properties, ecotoxicological effects and geo-referenced distribution patterns of uses, potential loads and potentially responding communities enables an integrated risk identification and assessment (see chapter 5.4.3).

5.4.2 Geo-referenced probabilistic risk assessment of pesticides - evaluation

Pesticide use according to Good Agricultural Practice (GAP) can contaminate surface waters. To date, the resulting risk for non-target organisms (e.g. algae, macrophytes, invertebrates and fish) is assessed by comparing the Predicted Environmental Concentration (PEC) with effect concentrations shown in ecotoxicological test. The PECs are assessed by using models calculating contamination of water bodies via spraydrift, drainage and runoff, assuming a set of realistic worst-case assumptions. The result is a specific PEC for a specific use and a defined distance between culture and water body (deterministic approach). By comparison of PECs for different distances with ecotoxicological effect concentrations, risk mitigation measures are requested, if necessary. These are PEC-mitigation measures, such as obligatory buffer zones of use of drift-reducing nozzles.

However, factors causing exposure of surface waters by pesticide use in the field differ locally and timely, in most situations probably resulting in lower exposure than calculated by multiple worst-case assumptions. The German UBA aims at the inclusion of landscape variability in pesticide risk identification by a probabilistic and geo-referenced approach to achieve more realistic risk assessments, more practicable risk mitigation requirements and thus enhanced acceptance as well as controllability. Nevertheless, it has to be ensured that simplified and/or facilitated exposure mitigation measures are still able to guarantee a sufficient level of protection. The approach includes the risk assessment for different landscapes on a local or regional scale and the identification of “hot spots”, which are water body segments with expected high risk. Risk management will be focussed on these high-risk segments, which can either be regionally grouped and defined as “Sondergebiet” (area with specific risk management) or managed locally.

The JKI institute for consequence assessment of plant protection elaborated the basic concept for a GIS-related probabilistic approach for drift exposure assessment in space cultures (fruits, wine, hops) (Golla et al. 2002). Based on this and on IVA-funded work by RLP Agroscience, UBA, BVL, JKI und IVA agreed on a conceptual framework for the realization of the approach in the execution of the German pesticide regulation[8]. An expert workshop held by the University of Landau (session chairs: Michael Klein, Fraunhofer IME, Udo Hommen Fraunhofer IME, Martin Bach, University of Gießen) evaluated and further developed the conceptual framework (Schulz et al. 2007) to comprise four steps:

1. National risk assessment including geo-referenced (geo-referenceable) factors
2. Hotspot analysis regarding the spatial extent and the severity of exposure as well as effect size tolerated by populations
3. Refined exposure assessment by using e.g. aerial photographs or field mapping
4. Risk management focusing on landscape-related exposure-mitigating measures that will reduce risk effectively and in a long-term perspective

The workshop identified several open points that have to be clarified scientifically before implementing the approach in the regulation. One focus is on setting criteria for the definition of hotspots, as the hotspot analysis is a new and crucial part of the assessment concept. As the new approach has to ensure crop protection as well as protection of the natural balance, ecological, regulatory and socio-economic consequences shall be evaluated in a scenario-based manner. A joint proposal by RLP Agroscience and my department at Fraunhofer IME won the competitive call. The R&D project funded by UBA (FKZ 3707 63 4001) and headed by Roland Kubiak, RLP Agroscience, and Udo Hommen, Ecotoxicology department, Fraunhofer IME was finanlized in 2011 (ref. 64). Important partners for exposure assessment were Michael Klein, Ecological chemistry department of Fraunhofer IME and Martin Bach, University of Gießen. For ecotoxicological considerations, RWTH Aachen and gaiac were included. The JKI contributes the basic database and GIS skills.

The consortium developed and evaluated a new approach for the aquatic risk assessment of plant protection products in Germany as part of the GeoRisk project. The aim was to establish a more realistic risk assessment procedure, allowing simplified and reduced substance-specific risk mitigation measures while maintaining the existing level of protection.

The GeoRisk approach is based on the following elements:

- Geodata-based probabilistic calculation of drift and volatilization / deposition entries and initial concentrations of plant protection products in edge-of-field water bodies adjacent to orchards, hops and grape cultures
- Consideration of the dispersion and transport of plant protection products in running waters using a dynamic exposure model
- Algorithm to create flow direction, volume and velocity based on ATKIS data (Trapp)

[8] UBA/BVL/BBA/IVA (2006) Probabilistische Bewertung des Umweltrisikos von Pflanzenschutzmitteln und Festsetzung von Anwendungsbestimmungen im Rahmen des Zulassungsverfahrens für Pflanzenschutzmittel. Konzept, Stand 06.02.2006.

- Identification of the ecologically critical aggregation of water body segments with high risks (“hotspots”) taking into account tolerable effect levels for populations of aquatic species
- Implementation of spatially differentiated generic risk management in the identified hotspots
- Authorization of products based on the risk of new product-related hotspots

One of the main outcomes of the project was a recommendation to use the more realistic dynamic exposure model instead of the formerly favoured static model. However, due to the complexity of the new model and missing data for several input parameters, it has not been possible during this project to achieve nationwide implementation. Therefore, the model has been applied to two representative streams in the hops region Hallertau. These preliminary results were extrapolated to all permanent crop areas in Germany (except for the “Altes Land”) which revealed approximately 200 km of anticipated management segments. We explored steps required to implement the approach, including the necessary hotspot management, in all permanent crop areas in Germany.

When applying the dynamic exposure model, situations with very high but very short concentrations peaks were simulated. For this type of exposure, ecotoxicological standard test results are inappropriate. We used carbaryl as a worst case substance (data see chapter 2.2.3). Due to its mode of action (inhibition of acytylcholinesterase) and very fast lethal effects on invertebrates it was not included in Annex I of 91/414/EC (now 1107/2009/EEC). In addition to LC50 values for exposure over 96 h, data were also available for 1 hour pulse exposure (followed by 95 h in untreated medium) for three sensitive insect taxa which indicate that an LC50 for 1 h exposure can be expected to be at least by a factor of 5 higher than for 96 h exposure. To obtain a relationship of LC50 from exposure duration, the EC50 over 24 and 48 h from another sensitive species, *D. longispina*, were also considered (Figure 56).

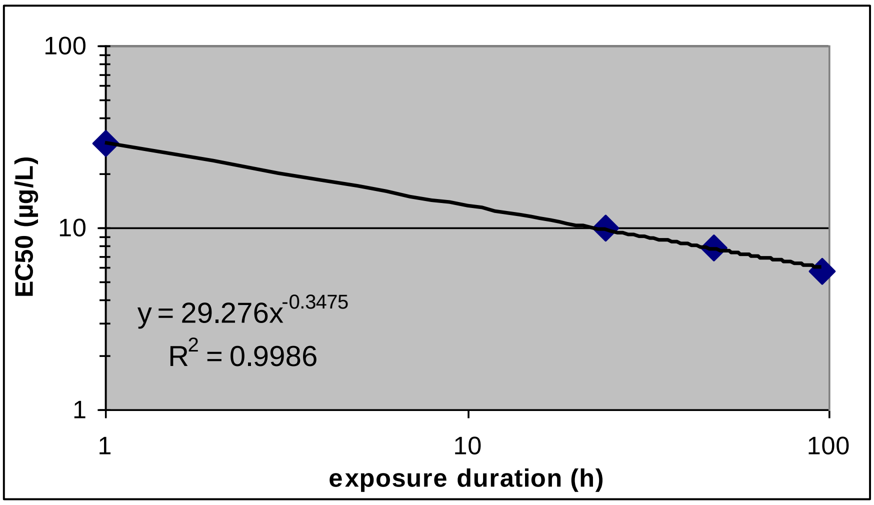


Figure 56: LC50 depending on the duration of exposure based on carbaryl data for the particularly sensitive species *Isoperla spec.* and *Daphnia longispina*

The relationship obtained for carbaryl was applied to exposure profiles predicted for a small stream in the intense hops culture area Hallertau, Germany. As a measure of exposure duration the time with PEC above the RAC (regulatory acceptable concentration based on ecotoxicity data with carbaryl) was used. While the comparison of the PECs with the standard RAC indicates high risk for most of the stream segments, the RAC* considering the shorter exposure duration was not exceeded in this example. The presented approach brings more realism into the risk assessment for situations characterized by short-term pulse exposure. Because the function estimating the LC50 for exposure duration shorter than in the standard test was based on the data of a very fast acting substance, the method should be protective for more slowly acting substances. A literature research did not reveal more sensitive relationships.

Before implementation of the approach, chronic effects after shortterm exposure have to be considered and handled, when necessary. The dynamic exposure model has to be applied to all small lotic water bodies in Germany. This will enable assessments of potential pesticide loadings of small tributary waterbodies when entering main rivers and produce inputs in georeferenced tools used in the water framework directive or for nature conservation. As the geodata for Germany span a wide range of topographical and hydrological conditions, it should be no problem to extend approach to the entire EU, which is not possible with e.g. the Dutch approach based on lowland drainage ditches only.

5.4.3 IRIS - Integrated Risk Identification and assessment of Substances

As contribution for a proposal of an excellence centre at the University of Gießen we developed the idea of IRIS, focusing all available data to be perceived and processed by a network (= retina) of all available experts to get a comprehensive view on substance-related risks. An integrated risk identification and assessment of environmental chemicals should include the protection goods ecosystem integrity and human health and account for all variants of potential risks.

To date, causal relationships necessary for the risk assessment are investigated within the respective scientific discipline, independently: from the playing together of the different contexts. The interaction intended by EU framework projects mostly resulted in a collection of aspects, which were hardly systematically connected and worked at interactively.

There is need for an interdisciplinary centre for the prediction, assessment and record of the fate and effects of environmental chemicals and their metabolites in ecosystems as well as in the food and feed chain or consumer products by integrative analytical methods and models. The trans-disciplinary concept connects different observation methods of chemical fate and effects by a central data acquisition and linkage to make them accessible for description and prediction models.

The structured data pool will serve to derive hypotheses for setting up and consequently process work programs. Finally, a risk assessment and risk-benefit analysis of chemical loads, specific applications and possible alternatives will be performed based on principles of sustainable economy. The data generation for the risk assessment by specific and integrative test systems will be triggered by the needs of identification, assessment and prediction of concentrations and effects. University experts for these fields will be linked and focussed by the needs of industry and authorities, which will be translated by the Fraunhofer IME. The technical college is included to develop and transfer innovative technologies for analysis and reduction of chemical emissions into the environment.

For the qualitative and quantitative analysis of substances, chemical analytical methods will be supplemented by effect-specific biological methods. Ecotoxicological effects will be assessed based on data by established and developing test systems on different levels of biological organization, and different possibilities of integration of organization levels. The assessments may be verified by data from monitoring programs and influence-related effect monitoring approaches (see chapter 3.4). Effects on human health are linked to actual organismic loads of individuals representing human sub-populations by evaluating human biomonitoring data of environmental medicin research programs by multivariate statistics.

The generation, adaption and application of models enable quantitative predictions (including variability) of the complex processes of behavior, fate and effects of substances.

- Emission models based on production, use patterns or market penetration factors and on the main emission pathways

- Quantitative structure-activity relationships (QSAR) for predicting environmentally relevant behavior and toxicological effects of chemicals from structural descriptors
- Process models for the spatial and temporal distribution of chemicals in the environment (long-range transport, distribution to different environmental compartments) and degradation, based on physico-chemical properties and degradation tests (hydrolysis, photolysis, biodegradation)
- Bioavailability models (e.g. biotic ligand models BLM)
- Uptake, distribution, depuration and metabolism models for bioconcentration and biomagnification processes, Modeling concentrations at effect sites
- Species sensitivity distribution models
- Population effect and recovery models of different complexity (individual, age-structure and recruitment models)
- Community models (to date rather simple and generic)
- Geographic information systems for integrating site-specific information of any type (see 5.4.2)

The spatial differentiation applies for different scales. Thus, risk assessments can be performed dependent on site-specific, landscape-specific, continentally or globally different environmental conditions. The organization of a substance-related data centre aiming at systematic collection of compatible data enables a worldwide view on chemical distribution and risk on different scales and hypotheses-driven evaluation for authorities and industry. For this, existing competences in the fields of bioinformatics, GIS, probabilistic quantification and dynamic modeling of fate and effects will interact and further develop.

Including geo-referenced information on nature conservation issues as well as socio-economics (see 5.4.2) links substance-related risk assessment with concepts of sustainable economic development. A centre comprising the multidisciplinary expertise and the necessary tools for modeling and geo-referencing is able to perform assessments for applications and uses of chemicals in any part of the world.

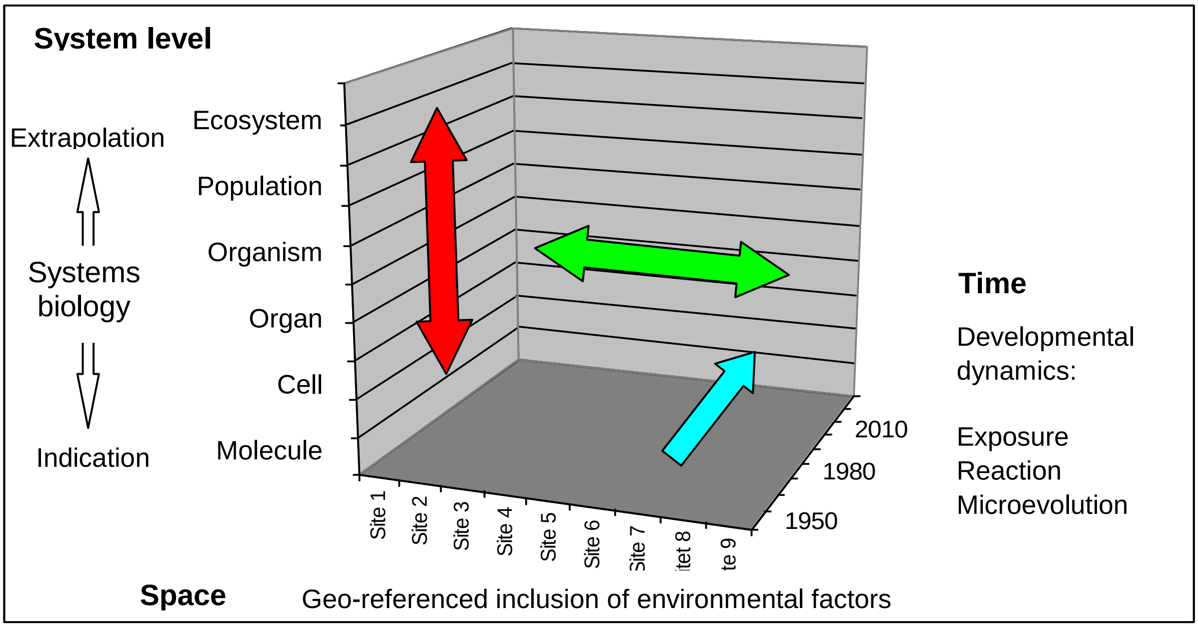


Figure 57: Dimensions to be integrated for a comprehensive risk assessment of environmental chemicals

6 References

Abbas R, Hayton WL (1997) A Physiologically Based Pharmacokinetic and Pharmacodynamic Model for Paraoxon in Rainbow Trout Toxicology and Applied Pharmacology, 145 (1): 192-201.

Allner B (1997) Toxikokinetik von 3,4-Dichloranilin beim dreistachligen Stichling (Gasterosteus aculeatus) unter besonderer Berücksichtigung der Fortpflanzungsphysiologie. Dissertation thesis, University of Mainz, 129 pages.

Bachmann J (2002) Entwicklung und Erprobung eines Teratogenitäts-Screening-Testes mit Embryonen des Zebrabärblings Danio rerio. Dissertation thesis, TU Dresden.

Beaty BJ, Black WC, Carlson,JO, Clements WH, DuTeau N, Harrahy E, Nuckols J, Olson K E, Rayms-Keller A (1998) Molecular and genetic ecotoxicological approaches to aquatic environmental bioreporting. Environmental Health Perspectives 106 (Supp. 6): 1395-1407.

Belanger SE, Farris JL, Cherry DS, Cairns J jr (1990) Validation of Corbicula fluminea growth reductions induced by copper in artificial streams and river systems. Canadian Journal of Fisheries and Aquatic Sciences 47(5): 904-914.

BfN (Bundesamt für Naturschutz, Bonn-Bad Godesberg) (ed) (1998) Rote Liste gefährdeter Tiere Deutschlands. Schriftenreihe für Landschaftspflege und Naturschutz 55: 168-230

BML (1998) Revision of the Act concerning the protection of crop plants (Plant Protection Act), 14. Mai 1998. BGBI. I: 971, 1527, 3512.

Borgmann U, Cove R, Loveridge C (1980) Effects of metals on the biomass production kinetics of freshwater copepods. Canadian Journal of Fisheries and Aquatic Sciences 37: 567-575.

Brock TCM, van Wijngaarden RPA, van Geest GJ (2000) Ecological risks of pesticides in freshwater ecosystems. Part 2: Insecticides. Alterra-Rapport 089. 142 pp., Alterra Green World Research, Wageningen, Netherlands.

Bundesgesetzblatt (2005) Jahrgang 2005 Teil I Nr. 5, ausgegeben zu Bonn am 25.01.2005. Bekanntmachung der Neufassung des Abwasserabgabengesetzes vom 18. Januar 2005.

Butler NM, Suttle CA, Neill WE (1989) Discrimination by freshwater zooplankton between single algal cells differing in nutritional status. Oecologia 78:368-72.

Calow PJ, Berryl RJ (eds) (1989) Evolution, ecology and environmental stress. Academic Press, London.

Campbell PJ, Arnold DJS, Brock TCM, Grandy NJ, Heger W, Heimbach F, Maund SJ, Streloke M (eds) (1999) Guidance document on higher-tier aquatic risk assessment for pesticides (HARAP). SETAC-Europe publication, Brussels.

Clements WH, Farris JL, Cherry DS, Cairns J jr (1989a) The influence of water quality on macroinvertebrate community responses to copper in outdoor experimental streams. Aquatic Toxicology 14: 249-262.

Clements WH, Cherry DS, Cairns J jr (1989b) The influence of copper exposure on predator-prey interactions in aquatic insect communities. Freshwater Biology 21: 483-488.

Cossland NO, Mitchell GC, Bennett D, Maxted J (1991) An outdoor artificial stream system designed for ecotoxicological studies. Ecotoxicology and Environmental Safety, 22: 175-183.

Dieterich B (2006) Entwicklung und Erprobung eines ökotoxikologischen Tests mit Makrophyten in 270 L-Aquarien. Diplome thesis, FH Bingen.

DIN 38415-6 (2001) German standard methods for the examination of water, waste water and sludge - subanimal testing (group T) - part 6: Toxicity to fish; Determination of the non-acute-poisonous effect of waste water to fish eggs by dilution limits (T 6). Deutsches Institut für Normung e.V.

Dooley K, Zon LI (2000) Zebrafish: a model system for the study of human disease. Current Opinion in Genetics & Development 10: 252-256.

Drenner RW, McComas SR (1980) The roles of zooplankter escape ability and fish size selectivity in the selective feeding and impact of planktivorous fish. In: Kerfoot WC (ed), Evolution and ecology of zooplankton communities, pp 587-93. University Press of New England, NE, USA.

Dülmer U (1998) Konkurrenzbeziehung zweier Cladocerenarten unter Fremdstoffeinfluss. Analyse anhand von Experimenten und individuenbasierter Simulation. Dissertation thesis, University of Aachen. Shaker-Verlag, Aachen - Maastricht.

Ensenbach U (1991) Kinetik und Dynamik von Fremdstoffgemischen beim Zebrabärbling (Brachydanio rerio). Dissertation thesis, Johannes Gutenberg-Universität Mainz

Ertz A (1996) Untersuchung der Biozönose eines künstlichen Fließgewässers unter dem Aspekt der Eignung als ökotoxikologisches Testsystem. Diploma Thesis, University of Aachen, Germany.

Golla B, Enzian S, Jüttersonke B, Gutsche V (2002) Entwicklung und Testung eines GIS-gestützten Verfahrens zur Erstellung thematischer Risikokarten als Grundlage für eine Differenzierung von Anwendungsbestimmungen zum Schutz des Naturhaushaltes beim Einsatz von Pflanzenschutzmitteln. Report 201 63 410; Umweltbundesamt Berlin.

Girling AE, Pascoe D, Janssen CR, Peither A, Wenzel A, Schäfer H, Neumeier B, Mitchell GC, Taylor EJ, Maund SJ, Lay JP, Jüttner I, Crossland NO, Stephenson RR, Persoone G (2000) Development of methods for evaluating toxicity to freshwater ecosystems. Ecotoxicology and Environmental Safety 45: 148-176.

Hammers-Wirtz M, Ratte HT (2000) Offspring fitness in Daphnia: Is the Daphnia reproduction test appropriate for extrapolating effects on the population level? Environ Toxicol Chem 19:1856-66.

Hanazato T, Hirokawa H (2004) Changes in vulnerability of Daphnia to an insecticide application depending on the population phase. Freshw Biol 49:402-9.

Hart BT, Currey NA, Jones MJ (1992) Biogeochemistry and effects of copper, manganese and zink added to enclosures in Island Billabong, Magela Creek, northern Australia. Hydrobiologia 230: 93-134.

Havens KE (1994a) An experimental comparison of the effects of two chemical stressors on a freshwater zooplankton assemblage. Environmental Pollution 84:245-251.

Havens KE (1994b) Structural and functional responses of a freshwater plankton community to acute copper stress. Environmental Pollution 86: 259-266.

Hedtke SF (1984) Structure and function of copper-stressed aquatic microcosms. Aquatic Toxicol. 5: 227-244.

Hund-Rinke K, Lindemann M, Simon M (2005) Experiences with novel approaches in earthworm testing alternatives. JSS – J Soils Seds 5:233-9.

Hutchinson TH, Ankley GT, Segner H, Tyler CR (2006) Screening and testing for endocrine disruption in fish – biomarkers as “signposts”, not “traffic lights”, in risk assessment. Environ. Health Perspectives 114: 106-114.

Jüttersonke, Barbara und Klaus Arlt (2004) Der Einfluss von Herbizidapplikationen auf Nichtzielpflanzen. Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz, Sonderheft XIX, 995-1002.

Kühne S, Freier B, Kaul P, Kalthoff N (2001) Einfluss von Pflanzenschutzmittels auf Heuschrecken (Saltatoria) in Saumstrukturen. Mitt. Dtsch. Ges. Allg. Angew. Ent. 13, 245-248

Länge R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP (2001) Effects of the synthetic estrogen 17 alpha-ethinylestradiol on the life-cycle of the fathead minnow (Pimephales promelas). Environ Toxicol Chem 20, 1216-27.

Leland HV, Carter JL (1984) Effects of copper on species composition of periphyton in a Sierra Nevada, California stream. Freshwater Biology 14: 281-296.

Leland HV, Carter JL (1985) Effects of copper on production of periphyton, nitrogen fixation and processing of leaf litter in a Sierra Nevada, California stream. Freshwater Biology 15: 155-173.

Leland HV, Kent E (1981) Effects of copper on microfaunal species composition in a Sierra Nevada, California stream. Verh. Internat. Verein. Limnol. 21: 819-829.

Leland HV, Fend SV, Dudley TL, Carter JL (1989) Effects of copper on species composition of benthic insects in a Sierra Nevada, California Stream. Freshwater Biology 21: 163-179.

Liess M, Schulz R (1999) Linking insecticide contamination and population response in an agricultural stream. Env Tox Chem 18, 9, 1948-1955

Liess M, Schulz R, Berenzen N, Nanko-Drees J, Wogram J (2001) Pflanzenschutzmittel-Belastung und Lebensgemeinschaften in Fließgewässern in landwirtschaftlich genutztem Umland (Pesticide contamination and macroinvertebrate communities in running waters in agricultural areas). UFP Bericht 29624511, UBA-FB 000187, UBA Berlin

Liess M, Brown C, Dohmen D, Duquesne S, Hart A, Heimbach F, Kreuger J, Lagadic L, Maund SJ, Reinert W, Streloke M, Tarazona JV (eds) (2005) Effects of pesticides in the field (EPIF). SETAC-Europe publication, Brussels.

Moore MV, Winner RW (1989) Relative sensitivity of Ceriodaphnia dubia laboratory tests and pond communities of zooplankton and benthos to chronic copper stress. Aquatic Toxicology 15: 311-330.

Müller A, Buhr L, Petemer W, Strumpf T (2003) Auswirkungen von FUNGURAN auf eine aquatische Lebensgemeinschaft sowie das Rückstandsverhalten von Kupfer in Wasser und Sediment. Gesunde Pflanzen 55 (9): 244-243.

OECD (2003) Draft test guideline on fish two-generation (March 2003 version). OECD Guidelines for the Testing of Chemicals, OECD, Paris, France.

OECD (2006a) Guidance document on simulated freshwater lenthic field tests (outdoor microcosms and mesocosms). OECD Guidelines for the Testing of Chemicals, OECD, Paris, France.

OECD (2006b) Fish Embryo Toxicity Test. Draft OECD guideline for the testing of chemicals. http://www.oecd.org/dataoecd/39/59/36817070.pdf.

OECD (2009) Test No. 230: 21-day Fish Assay. A short-term screening for oestrogenic and androgenic activity, and aromatase inhibition. OECD Guidelines for the Testing of Chemicals, Section 2: Effects on Biotic Systems, OECD Publishing, Paris, France.

OECD (2011) Test No. 234: Fish sexual development test. OECD Guidelines for the Testing of Chemicals, Section 2: Effects on Biotic Systems, OECD Publishing, Paris, France.

Nagel R (1986) Untersuchungen zur Eiproduktion beim Zebrabärbling (Brachydanio rerio, Ham.-Buch.). J. Appl. Ichtyol. 4, 173-181.

Nagel R (1998) Der Vollständige Life Cycle Test (Complete Life Cycle Test, CLC-Test) mit dem Zebrabärbling (Danio rerio, vormals Brachydanio rerio), Entwurf. UBA-Texte 58/98: 166-175.

Nagel R (2002) DarT: The embryotest with the zebrafish Danio rerio - a general model in ecotoxicology and toxicology. ALTEX 19 Suppl 1/02, 38-48.

Parlar H, Angerhöfer D (1995) Chemische Ökotoxikologie. Springer-Verlag Berlin, Heidelberg, New York, 386 pages

Parrot JL, Blunt B, Sullivan CV, Rhodes S (2005) Endocrine disrupting effects in fathead minnows exposed for a lifecycle to pharmaceuticals and municipal wastewater effluent. Poster Presentation at the 15th Annual Meeting of SETAC Europe, Lille, 22.-26. Mai 2005.

Pawlowski S, van Aerle R, Tyler CR, Braunbeck T (2004) Effects of 17alpha-ethinylestradiol in a fathead minnow (Pimephales promelas) gonadal recrudescence assay. Ecotoxicology and Enviromental Safety, 57, 302-317.

Preuß T (2006) Ökotoxikologische Charakterisierung von Nonylphenol Isomeren. Dissertation thesis, RWTH Aachen

Reh W (1991) Populationsbiologische Untersuchungen am Grasfrosch (Rana temporaria L., 1758). Dissertation thesis, University of Mainz. Dissertations Druck Darmstadt.

Reh W, Seitz A (1990) The influence of land use on the genetic structure of populations of the common frog Rana temporaria. Biological conservation 54, 239-249.

Roß-Nickoll M, Lennartz G, Fürste A, Mause R, Ottermanns R, Schäfer S, Smolis M, Theißen B, Toschki A, Ratte HAT (2004) Die Arthropodenfauna von Nichtzielflächen und die Konsequenzen für die Bewertung der Auswirkungen von Pflanzenschutzmitteln auf den terrestrischen Bereich des Naturhaushaltes. Bericht. Im Auftrag des Umweltbundesamtes, FKZ 200 63 403

Roussell H (2005) Les effects du cuivre sur la structure et le fonctionnement des ecosystems aquatiques: Une étude en mesocosmes lotiques. Effects of copper on structure and function of freshwater ecosystems: a lotic mesocosm study. Dissertation thesis, University of Toulouse III, France (C. Schäfers was involved in the study set-up discussions and acted as external rapporteur).

Rudolph P, Boje R (1986) Ökotoxikologie. Grundlagen für die ökotoxikologische Bewertung von Umweltchemikalien nach dem Chemikaliengesetz. Ecomed Fachverlag, Landsberg.

SANCO (2002) Guidance Document on Aquatic Ecotoxicology in the context of the Directive 91/414/EEC. Working Document. European Commission, Health & Consumer Protection Directorate-General. Sanco/3268/2001 rev.4 (final). 17 October 2002.

Scanferlato VS, Cairns J jr (1990) Effect of sediment-associated copper on ecological structure and function of aquatic microcosms. Aquatic Toxicology 18: 23-34.

Scholz S, Kordes C, Hamann J, Gutzeit HO (2004) Induction of vitellogenin in vivo and in vitro in the model teleost medaka (Oryzias latipes): comparison of gene expression and protein levels. Mar Environ Res 57, 235-44.

Schulz R, Elsaesser D, Ohliger R, Stehle S, Zenker K (2007) Umsetzung der georeferenzierten probabilistischen Risikobewertung in den Vollzug des PflSchG- Pilotphase – Dauerkulturen. FKZ 206 63 402, Umweltbundesamt, Dessau.

Seitz A., Loeschke V (1991) Species conservation: A population-biological approach. Birkhäuser Verlag, Basel.

Sheahan DA, Bucke D, Matthiessen P, Sumpter JP, Kirby MF, Neall P, Waldock M (1994) The effects of low levels of 17 alpha ethynylestradiol upon plasma vitellogenin levels in male and female rainbow trout, Oncorhynchus mykiss, held at two acclimation temperatures. Presentation at the 17th session of EIFAC 1992, 99-112.

Takahashi H (1977) Juvenile Hermaphroditism in the zebrafish, Brachydanio rerio. Bull. Fac. Fish. Hokkaido Univ. 28 (2), 57-65.

Taub FB, Kindig AC, Meador JP, Swartzmann GL (1991) Effects of 'seasonal succession' and grazing on copper toxicity in aquatic microcosms. Verh. Internat. Verein. Limnol. 24: 2205-2214.

Tristsch H, Freier B, Kühne S, Karg W, Stark A (2001) Wiederbesiedlung und Erholung von lokalen Arthropodenpopulationen in Feldsäumen nach einem chemischen Knock-out. Mitt. Dtsch. Ges. Angew. Ent. 13, 237-240Winner RW, Owen HA, Moore MV (1990) Seasonal variability in the sensitivity of freshwater lentic communities to a chronic copper stress. Aquatic Toxicology 17: 75-92.

Van Den Brink PJ, Ter Braak CJF (1999) Principal Response Curves: Analysis of time-dependent multivariate responses of a biological community to stress. Environmental Toxicology and Chemistry 18:138-148.

Van den Brink PJ, Roelsma J, Van Nes EH, Scheffer M, Brock TCM (2002) PERPEST model, a case-based reasoning approach to predict ecological risks of pesticides. Environmental Toxicology And Chemistry. 21: 2500-2506.

van Wijngaarden RPA, Cuppen JGM, Arts GHP, Crum SJH, van den Hoorn MW, van den Brink PJ, Brock TCM (2004) Aquatic risk assessmnet of a realistic exposure to pesticides used in bulb crops: a microcosm study. Env. Tox. Chem .23: 1479-1498

Wilson EO, Bossert WH (1973) Einführung in die Populationsbiologie. Heidelberger Taschenbücher 133, Springer-Verlag, Berlin, Heidelberg, 168 p. Original: Wilson EO, Bossert WH (1971) A Primer of Population Biology.

Winfield IJ, Peirson G, Cryer M, Townsend CR (1983) The behavioural basis of prey selection by underyearling bream (Abramis brama (L.)) and roach (Rutilus rutilus (L.)). Freshw Biol 13: 139-49.

Winner RW, Owen HA (1991) Seasonal variability in the sensitivity of freshwater phytoplankton communities to a chronic copper stress. Aquatic Toxicology 19: 73-88.

Winner RW, Van Dyke JS, Caris N, Farrel MP (1975) Response of the macroinvertebrate fauna to a copper gradient in an experimentally polluted stream. Verh. Internat. Verein. Limnol. 19, 2121-2127.

Wright JF, Sutcliffe DW, Furse MT (eds) (2000) Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Young W, Whitehouse P, Johnson I, Brighty G (2002) Steroid oestrogens: Environmental thresholds for the protection of freshwater life. WRc-NSF R&D Technical Report P2-T04/1. Also: Poster at 12th Annual Meeting SETAC Europe. Wien 12.-16. Mai 2002.

7 Own Publications (in the order of publication)

7.1 Theses

1) Schäfers C (1988) Untersuchungen zum Reproduktionsverhalten und zur Populationsdynamik des Zebrabärblings (*Brachydanio rerio* Ham.-Buch.) in einem naturnahen Laborsystem. (*Investigations on behaviour, reproduction and population dynamics of zebrafish (Danio rerio) in a laboratory steam system*) Diplome thesis, Fachbereich Zoologie, Johannes-Gutenberg-University Mainz, Germany, 110 pages.

2) Schäfers C (1991) Toxizität und Populationsökologie - Einfluß von 3,4-Dichloranilin auf Fische mit unterschiedlichen Reproduktionsstrategien. (*Toxicity and population ecology – effects of 3,4-dichloroaniline on fish with different strategies of reproduction*.) Dissertation thesis, Fachbereich Zoologie, Johannes-Gutenberg-University Mainz, Germany, 174 pages. Since 2008 available as ISBN 978-3-8167-7505-8, Fraunhofer IRB Verlag.

3) Teigeler M (2012) Biomarker für sexual-endokrine Wirkungen beim Zebrabärbling (Danio rerio) Vergleich zwischen Kurzzeit-Belastung und Daten aus Full Life Cycle Tests (*Comparison and evaluation of biomarker responses in short-term and full life cycle tests with zebrafish*). Dissertation thesis, Technical Universiy of Aachen, Germany.

7.2 Peer reviewed papers and articles

4) Schäfers C, Nagel R, Seitz A (1989) Verhalten, Reproduktion und Populationsdynamik des Zebrabärblings (Brachydanio rerio Ham.-Buch.) in einem naturnahen Laborsystem. (*Behaviour, reproduction and population dynamics of zebrafish in a laboratory stream system*) Fischökologie 1 (2): 45-59.
Part of the diplome thesis.

5) Schäfers C, Nagel R (1991) Effects of 3,4-dichloroaniline on fish populations. Comparison between r- and K-strategists: A complete life cycle test with the guppy (Poecilia reticulata). Archives of Environmental Contamination and Toxicology 21: 297-302.
Part of the dissertation thesis.

6) Schäfers C, Nagel R (1993) Toxicity of 3,4-dichloroaniline to perch (Perca fluviatilis) in acute and early life stage exposures. Chemosphere, 26 (9): 1641-1651.
Part of the dissertation thesis.

7) Schäfers C, Nagel R (1993) Consequences of 3,4-dichloroaniline to guppy populations (Poecilia reticulata): computer simulation and experimental validation. The Science of the Total Environment, Supp. 2: 1471-1478.

8) Schäfers C, Oertel D, Nagel R (1993) Effects of 3,4-dichloroaniline on fish populations with differing strategies of reproduction. In: Braunbeck T, Hanke W, Segner H (eds) Fish. Ecotoxicology and ecophysiology. VCH Verlagsgesellschaft, Weinheim, p. 133-146.

9) Schäfers C, Nagel R (1994) Fish toxicity and population dynamics: effects of 3,4-dichloroaniline and the problems of extrapolation. In: Müller R, Lloyd R (eds) Sublethal and chronic effects of pollutants on freshwater fish. FAO and fishing new books, Blackwell Scientific Publishers, Oxford, p. 229-238.

10) Debus R, Fliedner A, Schäfers C (1996) An artificial stream mesocosm to simulate fate and effects of chemicals: technical data and initial experience with the biocenosis. Chemosphere 32 (9): 1813-1822.
Part of the post-doc position work: The technical system had been patented by R. Debus, head of the ecotoxicology department of the Fh-IUCT. A. Fliedner was the head of the aquatic ecotoxicology lab. The entire installation of the community including selection of the reference brook, the generation of biological data and the manuscript preparation was done by C. Schäfers.

11) Fliedner A, Remde A, Niemann R, Schäfers C (1997) Effects of the organotin pesticide azocyclotin in aquatic microcosms. Chemosphere 35 (1/2), 209-222.
C. Schäfers was responsible for the determination of the planktonic algae and contributed to the ecological interpretation.

12) Schäfers C, Klein W (1999) Ökologische Ansätze in der Ökotoxikologie als Herausforderung für die Risikobewertung. In: Oehlmann J, Markert B (eds) Ökotoxikologie – ökosystemare Ansätze und Methoden. Ecomed Verlagsgesellschaft AG & Co. KG, Landsberg, 49-62.

13) Schäfers C, Hassink J (2000) Experimental simulation of fate in rivers: A mesocosm for studies with radiolabelled substances. In: Cornejo J, Jamet P (eds) Pesticide/soil interactions. Some current research methods. INRA editions, Paris: 389-395.

14) Segner H, Caroll K, Fenske M, Janssen CR, Maack G, Pascoe D, Schäfers C, Vandenbergh GF, Watts M, Wenzel A (2003) Identification of endocrine-disrupting effects in aquatic vertebrates and invertebrates : report from the European IDEA project. Ecotoxicology and Environmental Safety 54: 302-314.
C. Schäfers was responsible for the fish life cycle studies.

15) Segner H, Navas JM, Schäfers C, Wenzel A (2003) Potencies of estrogenic compounds in in vitro screening assays and in life cycle tests with zebrafish in vivo. Ecotoxicology and Environmental Safety 54 :315-322.
C. Schäfers was responsible for the fish life cycle studies.

16) Fenske M, Maack G, Schäfers C, Segner H (2005) An environmentally relevant concentration of estrogen induces arrest of male gonad development in zebrafish, Danio rerio. Environmnetal Toxicology and chemistry 24 (5): 1088-1098.
C. Schäfers was responsible for the fish life cycle study and the main result.

17) Schäfers C, Hommen U, Dembinski M, Gonzalez-Valero JF (2006) Aquatic macroinvertebrates in the Altes Land, an intensely used orchard region in Germany. Correlation between community structure and potential for pesticide exposure. Environmental Toxicology and Chemistry 25: 3275-3288.

18) Schäfers C, Teigeler M, Wenzel A, Maack G, Fenske M, Segner H (2006) Concentration- and time-dependent effects of the synthetic estrogen, 17α-ethynylestradiol, on reproductive capabilities of the zebrafish, *Danio rerio*. Journal of Toxicology and Environmental Health, Part A, 70: 768-779.

19) Schäfers C, Klöppel H, Takahashi Y (2007) Zooplankton avoidance behavior following spray drift exposure to a respiratory chain inhibitor. Human and Ecological Risk Assessment 13 (3): 527-534.

20) Fliedner A, Schäfers C (2007) Wassergefährdungspotenzial nativer Öle und Fette: Berücksichtigung physikalischer Effekte *(Freshwater hazard potential of native oils and fats: consideration of physical effects)*. UWSF– Z Umwelchem Ökotox 19 (2) 103-107.

21) Rossteuscher S, Schmidt-Posthaus H, Schäfers C, Teigeler M, Segner H (2007) Background pathology of the ovary in a laboratory population of zebrafish (Danio rerio). Disease of aquatic Organisms (DAO), 79 (2):169-72.

22) Belanger SE, Sanderson H, Fisk PR, Schäfers C, Mudge SM, Willing A, Kasai Y, Nielsen AM, Dyer SD, Toy R (2008) Assessment of the Environmental Risk of Long Chain Aliphatic Alcohols. . Ecotoxicology and Environmental safety, 72 (4): 1006-1015.

23) Schäfers C, Boshof U, Jürling H, Belanger SE, Sanderson H, Dyer SD, Nielsen AM, Willing A, Gamon K, Kasai Y, Eadsforth CV, Fisk PA, Girling AE (2008) Environmental properties of long chain alcohols. Part 2: Structure-activity relationship for chronic aquatic toxicity of long chain alcohols. Ecotoxicology and Environmental Safety, 72 (4): 996–1005.

24) Sanderson H, Belanger SE, Fisk PR, Schäfers C, Veenstra G, Nielsen AM, Kasai Y, Willing A, Dyer SD, Stanton K, Sedlak R (2008) An Overview of Hazard and Risk Assessment of the OECD High Production Volume Chemical Category – Long Chained Alcohols [C6-C22] (LCOH) Ecotoxicology and Environmental safety, 72 (4): 973-979.

25) Schäfers C, Frische T, Stolzenberg H.-C., Weyers A, Zok S, Knacker T (2008) Zur Entwicklung einer Prüfstrategie auf sexual-endokrine Wirksamkeit einer chemischen Substanz bei Fischen (*a testing strategy for sexual-endocrine effects of chemicals in fish)*. Umweltwissenschaften und Schadstoff-Forschung 20: 229-233.

26) Crane M, Groß M, Matthiessen P, Ankley G, Axford S, Bjerregaarde P, Brown R, Chapman P, Dorgeloh M, Galay-Burgosi M, Green J, Hazlerigg C, Janssen J, Lorenzen K, Parrott J, Rufli H, Schäfers C, Seki M, Stolzenberg HC, van der Hoeven N, Vethaak D, Winfield I, Zok S, Wheeler J (2010) Multi-Criteria Decision Analysis of Test Endpoints for Detecting Endocrine Active Substances in Fish Full Life Cycle Tests. Integrated Environmental Assessment and Management, 6 (3): 378-389.

27) Knacker T, Boettcher M, Rufli H, Frische T, Stolzenberg HC, Teigeler M, Zok S, Braunbeck T, Schäfers C (2010) Environmental Effect Assessment for Sexual Endocrine Disrupting Chemicals – Fish Testing strategy. Integrated Environmental Assessment and Management, 6 (4): 653-662.

7.3 Monographs

28) Holler S, Schäfers C, Sonnenberg J (1996) Umweltanalytik und Ökotoxikologie. Springer Verlag, Berlin, Heidelberg, 478 S.

J. Sonnenberg was the employer and contributed the sub-chapter on chromatographic methods (15 pages). S. Holler prepared the chapters 2 (environmental fate of chemicals, 124 pages) and 3 (analysis of environmental samples, 40 pages). C. Schäfers prepared the introduction and the chapters 1 (integrative levels of biological organisation, 145 pages) and 4 (interactions between substances and biological systems, 110 pages).

29) Schäfers C (1999) Darstellung und vergleichende Bewertung nationaler und internationaler Ansätze zur Klassifizierung der Beschaffenheit von Fließgewässern. Executive summary in English: Description and comparative assessment of national and international approaches to the classification of river health. UBA-Texte 21/99, ISSN 0722-186X, 202 pages.

30) Schäfers C, Egert E, Luckow T, Sehr I, Wenzel A (2001) Ökotoxikologische Prüfung von Pflanzenschutzmitteln hinsichtlich ihres Potentials zur Grundwassergefährdung. Executive summary in English: Ecotoxicological testing of pesticides with respect to their potential of endangering groundwater communities. UBA-Texte 76/01, ISSN 0722-186X, 87 pages.

31) Giddings J, Heger W, Brock TCM, Heimbach F, Maund S, Norman S, Ratte HT, Schäfers C, Streloke M (Eds) (2002) Community Level Aquatic System Studies Interpretation Criteria. Proceedings from the CLASSIC workshop held at the Fraunhofer Institute Schmallenberg, Germany, 30.05.-02.06.1999 (organized by C. Schäfers). Setac Press, 44 pages.

C. Schäfers initiated the workshop, was OC member, rapporteur of the OC meetings, local organisator and finally responsible for the funding.

32) Schäfers C, Halbur M (2002) Ideenwettbewerb Risikominimierungsmaßnahmen zum Schutz des Naturhaushaltes vor schädlichen Auswirkungen durch Pflanzenschutzmittel. Executive summary in English: Idea competition – measures for the mitigation of risk to the economy of nature by adverse effects of crop protection products. UBA-Texte 46/02, ISSN 0722-186X, 86 pages.

7.4 Invited papers; articles in books or conference proceedings

33) Oertel D, Schäfers C, Poethke HJ, Nagel R, Seitz A (1991) Simulation der Populationsdynamik des Zebrabärblings in einem naturnahen Laborsystem. Verhandlungen der Gesellschaft für Ökologie 20 (2): 865-869.

34) Schäfers C, Nagel R (1994) Fische in der Ökotoxikologie: Toxikologische Modelle und ihre ökologische Relevanz. Biologie in unserer Zeit 24 (4): 185-191.

35) Hassink J, Schäfers C, Fliedner A, Kördel W, Debus R (1996) Verbleib eines Pflanzenschutzmittels in einem künstlichen Freiland-Fließgewässer. Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft Berlin-Dahlem 321 (1996), p. 378.

36) Schäfers C, Hassink J (1996) Experimental simulation of pesticide loading in a stream mesocosm. Proceedings 1996 Brighton Crop Protection Conference - Pest and Diseases, p. 97-102.

37) Schäfers C (1998) Die Bedeutung endokriner Wirkungen von Fremdstoffen für Fischpopulationen. Tagungsbericht der DGL/SIL in Frankfurt am Main, 22.-26.09.1997, p. 911-915.

38) Wenzel A, Schäfers C (1998) Possibilities and limitations of standardised ecotoxicological tests – use of other ecologically relevant endpoints? Report of the International (EU) Workshop „Environmental risk assessment for veterinary medicinal products – scientific issues in the implementation of EMEA/CV/055/96“ on 17.-18.11.1997 (ISBN 3-00-003675-X), p. 153-160

39) Maak G, Fenske M, Schäfers C, Schmitz A, Segner H (1999) Development of zebrafish exposed to low levels of ethinylestradiol during different life stages. J. Fish Biol 55, Suppl A; 245-246.

40) Schäfers C., Wenzel A, Maack G (2001) Ethinylestradiol – a contraceptive for fish? Abstract. Archives of Pharmacology, Supplement to Vol 363 (4): R 170.

41) Schäfers C (2001) The impact of plant protection products on aquatic species. In: Kuhl M, Schmitz PM, Wiegand S (eds) Cost-Benefit-Analysis of Crop Protection. Wissenschaftsverlag Vauk Kiel KG; p. 124-132.

42) Schäfers C, Dembinski M, Jahn W (2001) Effect of the use of plant protection products on aquatic species – a monitoring approach. In: Forster R, Streloke M (eds) Workshop on Risk Assessment and Risk Mitigation Measures in the Context of the Authorization of Plant Protection Products (WORMM), Mitteilungen aus der Biologischen Bundesanstalt für Land- und Forstwirtschaft Berlin-Dahlem, Heft 383: 57-62.

43) Schmitz A, Schäfers C, Troß R, Ruch B, Kleeberg H (2001) Effects of NeemAzal on fish : full life cycle test with the zebrafish, Danio rerio. In : Reginerio S, Kleeberg H (eds) Practice oriented results on use and production of plan extracts and pheromones in integrated and biological pest control. Proceedings of the 2. Workshop « Neem & Pheromones » May 15-16, 2001, University of Uberaba, p. 37-50.

44) Schäfers C, Schmitz A, Wenzel A (2002) Effects of xenoestrogens on the fish life cycle. Proceedings of the 2nd BMBF/BMU/UBA Statusseminar „Endocrine Disruptors“, Berlin, 02.-04.04.2001, p. 97-98.

45) Schäfers C (2003) Auswirkungen von Arzneimitteln und endokrin wirksamen Substanzen auf die aquatische Lebensgemeinschaft. In : Track T, Kreysa (eds) Spurenstoffe in Gewässern. WILEY-VCH VerlagGmbH & Co KGaA, Weinheim, p. 47-60.

46) Schäfers C (2004) Kleine Zebras schwimmen im Dienste des Gewässerschutzes. In: Bullinger H-J (Eds.) Trendbarometer Technik. Visionäre Produkte, Neue Werkstoffe, Fabriken der Zukunft. Carl Hanser Verlag München, Wien, p. 252-253.

47) Schäfers C, Teigeler M, Wenzel A (2004) Endokrin wirksame Stoffe in der Umwelt: Die Anti-Baby-Pille für Fische. GIT Laborfachzeitschrift 48 (4): 330-333.

48) Schäfers C, Hommen U, Dembinski M, Gonzalez-Valero JF (2005) Landscape level pesticide risk assessment using freshwater macroinvertebrate community structure in the orchard agriculture of “Altes Land”, Germany. In: Liess M, Brown C, Dohmen P, Duquesne S, Hart A, Heimbach F, Kreuger J, Lagadic L, Maund S, Reinert W, Streloke M, Tarazona JV (eds) Effects of pesticides in the field. SETAC Press, p 75-77.

49) Schäfers C, Teigeler M (2005) Testen und Bewerten – eine Teststrategie für Fische (Testing and assessment – a test strategy for fish). In: Umweltbundesamt (ed.) 3. Statusseminar “Chemikalien in der Umwelt mit Wirkung auf das endocrine System. Wissenschaftliche Grundlagen der Bewertung und Regulierung (3rd UBA-Statusseminar on endocrine disrupting chemicals in the environment. Scientific basis of assessment and regulation), Fraunhofer IRB Verlag, Munich (ISBN 3-8167-6968-3), p 105-115.

50) Schäfers C (2007) Grundproblematik gefährlicher Stoffe in Gewässern. Gewässerschutz * Wasser * Abwasser 207: 32/1 - 32/8.

51) Schäfers C (2009) Umweltrisiko industrieller Wasserbausteine - Vorstellung eines Untersuchungs-konzepts. Bundesanstalt für Gewässerkunde, Veranstaltungen 5/2008: 107-119.

7.5 Non-GLP-reports not covered by publications

52) Kördel W, Hund K, Schäfers C, Schmitz A, Fliedner A, Hassink J (1998) Verbleib und Wirkung von Pflanzenschutzmitteln in Oberflächengewässern – Aquatische Mikrokosmosuntersuchungen. (Fate and effects of pesticides in surface waters – aquatic microcosm studies). Report FKZ 126 05 111, Umweltbundesamt Berlin, 86 pages. Appendix 1: Tables and Figures. Appendix 2: Guidance.

C. Schäfers prepared the final version and contributed appendix 2.

53) Wenzel A, Schäfers C (2001) Research efforts towards the development and validation of a test method for the identification of endocrine disrupting chemicals. Report to EU DG 24, B6-7920/98/000015, 80 pages.

C. Schäfers provided the fish life cycle studies and the conceptual conclusions.

54) Schäfers C, Delbeke K (2003) Community level study with copper sulphate, permanent exposure, in aquatic microcosms. Sponsor: International Copper Association (ICA), represented by the European Copper Institute (ECI). Final report, 102 pages.

55) Hommen U, Schäfers C, Roß-Nickoll M, Ratte HT (2004) Auswertung der wichtigsten in Deutschland durchgeführten Monitoring-Studien zur Auswirkung von Pflanzenschutzmitteln auf Nichtzielorganismen. BVL, 95 pages.

C. Schäfers contributed to the concept and the final conclusions.

56) Schäfers C, Teigeler M (2004) Zebrafish full life cycle tests: unaffected fish data. Document prepared for the OECD Fish Drafting Group, 14 pages.

57) Nagel R, Luwichowski KU, Oetken M, Schmidt J, Jackson P, Petersen G, Meller M, Eberling M, Tomczyk S, Schallnass HJ, Knacker T, Boshof U, Teigeler M, Schäfers C, Schwaiger J, Triebskorn R (2004) Ringtest zur Validierung der Prüfrichtlinie Fish Life-Cycle Test mit dem Zebrabärbling (Danio rerio). Forschungsbericht FKZ 200 67 411, Umweltbundesamt Berlin, 141 pages.

C. Schäfers was responsible for the final study plan and the VTG measurements.

58) Schäfers C, Teigeler M (2005) Validierung des 21-Tage Fischtests zur Erfassung von endokrin-wirksamen Stoffen. Forschungsbericht FKZ 203 65 420/02, Umweltbundesamt Berlin, 42 pages.

59) Schäfers C (2005) Abschätzung möglicher Auswirkungen von Kupferkonzentrationen, die in Gräben des Alten Landes gemessen wurden (*assessment of hazard potential of copper concentrations measured in ditches of the „Altes Land"*). Expert opinion prepared for the Federal German Agency for Consumer protection and food safety, BVL, 26 pages.

60) Schäfers C (2007) Abschätzung der Sicherheit einer Extrapolation von Wachstumsdaten aus Early Life Stage und Juvenile Growth Tests (OECD 210, 204, 215) auf die NOEC von Fish Full Life Cycle Tests bei der Risikobewertung von DMI-Fungiziden (*Assessment of the uncertainty of an extrapolation to a fish full life test NOEC from growth data from fish early life stage and juvenile growth tests (OECD 210, 204, 215) for the aquatic risk assessment of DMI fungicides*). Expert opinion prepared for the German agrochemical association IVA, 16 pages.

61) Schäfers C, Teigeler M, Knacker T (2007) Charakterisation of endocrine effects in fish: relevant parameters for the development of a new OECD test guideline and the use in the legislatory environmental risk assessment (German, abstract in English). Forschungsbericht FKZ 206 67 470, Umweltbundesamt Dessau, 97 pages.

62) Braunbeck T, Schäfers C, Teigeler M, Knacker T (2008) Studies to support the validation of a definitive fish test (full life-cycle / two-generation) (German, abstract in English). Forschungsbericht FKZ 204 67 454/02, Umweltbundesamt Dessau.

63) Schäfers C (2010) Potential criteria for exclusion of active substances in plant protection products from registration due to endocrine effects under the new regulation 1107/2009/EC - Expert opinion prepared for the Federal German Agency for Consumer protection and food safety, BVL, 29 pages.

64) Kubiak R, Hommen U, Bach M, Classen S, Golla B, Guerniche D, Klein M, Krumpe J, Preuss TG, Ratte HT, Roß-Nickoll M, Schäfers C, Strauss T, Toschki A, Trapp M (2010) GeoRisk - Geodata based Probabilistic Risk Assessment of Plant Protection Products (Georeferenzierte Probabilistische Risikobewertung von Pflanzenschutzmitteln). Report. UBA Project code 3707 63 4001, Umweltbundesamt Dessau.

7.6 Selected GLP higher tier study reports

SSD

65) Schäfers C. GLP-codes ACS-001/4-26/H, /I, /J, /K, /L, /M (2002) Acute invertebrate toxicity tests with sediment. Species: *Daphnia longispina* (H), *Chydorus sphaericus* (I), *Planorbarius corneus* (J), *Sphaerium corneum* (K), *Gammarus fossarum* (L), *Ephemera danica* (M).

Schäfers C. GLP-codes ACS-001/4-26/Na, /Nb (2002) Acute toxicity to *Chloroperla grammatica*, 96 h exposure (Na), 1 h exposure (Nb).

Schäfers C (2002) Carbaryl. Hazard potential to freshwater invertebrates. Proposed Ecologically Acceptable Concentration. Expert opinion.

Schäfers C, Hommen U (2004) Addendum to the expert opinion: Carbaryl. Hazard potential to freshwater invertebrates. Proposed Ecologically Acceptable Concentration.

66) Schäfers C. GLP-code SIM-002/4-80/A, /B, /C, /D, /E, /F, /G (2005) Single species toxicity tests on the effects of a herbicide. Species: *Myriophyllum spicatum* (A), *Ceratophyllum demersum* (B), *Elodea densa* (C), *Chara intermedia* (D), *Heteranthera zosterifolia* (E), *Hygrophila polysperma* (F), *Vallisneria spiralis* (G).

Hommen U, Schäfers C (2005) Probabilistic assessment of effects of a herbicide on aquatic macrophytes. Expert opinion.

Microcosm/mesocosm studies

67) Schäfers C. GLP-code URA-001/4-50 (2000) Community level study with a copper-based fungicide in aquatic microcosms. 123 pages + CD. Sponsor: European copper task force. In addition: Bridging-studies with six different products.

68) Schäfers C. GLP-code CYA-020/4-50 (2001) Community level effects of a herbicide in an indoor semi-realistic microcosm study. 76 pages + CD.

69) Schäfers C. GLP-code ACS-002/4-50 (2002) Zooplankton community level study with an acetylcholin-esterase inhibitor (insecticide) in semi-realistic aquatic microcosms. 105 pages + CD.

70) Wellmann P, Schäfers C. GLP-code MAR-001/4-50 (2003) Zooplankton community level study with an acetylcholin-esterase inhibitor (insecticide) in semi-realistic aquatic microcosms. 65 pages + appendices A-E. 119 pages + CD.

71) Wellmann P, Schäfers C. GLP-code NEU-001/4-50 (2004) Aquatic invertebrate study on the effect of an oil in indoor semi-realistic microcosms. 110 pages.

72) Schäfers C, Hommen U, Ebke P. GLP-code VIS-001/4-52 (2005) Aquatic mesocosm study on the effects of a chlororganic pesticide in an enclosure system in Homberg/Ohm. 158 pages + CD.

73) Schäfers C, Hommen U, Ebke KP Klöppel H. GLP-code FCS-011/4-52 (2006) Aquatic mesocosm study on the effects of a herbicide product in an enclosure system in Homberg/Ohm. 125 pages + Appendices A-K.

74) Schäfers C, Hommen U, Ebke KP, Klöppel H. GLP-code FCS-012/4-52 (2006) Aquatic mesocosm study on the effects of a herbicide product in an enclosure system in Homberg/Ohm. 133 pages + appendices A-L.

75) Wellmann P, Schäfers C, Hommen U. GLP-code FCS-010/4-52 (2006) Aquatic mesocosm study on the effects of a fungicide in an outdoor multiple pond system in Aachen. 136 pages + appendices A-L.

76) Wellmann P, Schäfers C, Müller J. GLP-code SIN-001/4-50 (2006) Aquatic invertebrate study on the effect of a paraffin oil in indoor semi-realistic microcosms. 110 pages.

77) Hommen U, Schäfers C, Böhmer W. GLP-code SCC-001/4-50 (2007) Zooplankton community level study with a respiratory chain inhibitor (acaricide) in semi-realistic aquatic microcosms. 168 pages + CD.

78) Schäfers C, Hommen U, Böhmer W, Strauss T, Ebke KP. GLP-code SIN-003/4-52 (2007) Plankton community study with a herbicide product in aquatic mesosocms . 102 pages + appendices 1-5.

79) Hommen U, Schäfers C, Ebke KP, Böhmer W. GLP-code FEI-013/4-52 (2007) Aquatic mesocosm study on the effects of a herbicide product in an enclosure system in Homberg/Ohm. 116 pages + appendices A-M.

80) Hommen U, Schäfers C, Ebke KP, Klöppel H. GLP-code FEI-021/4-52 (2008) Aquatic mesocosm study on the effects of a herbicide product in an enclosure system in Homberg/Ohm. 116 pages + appendices A-N.

81) Hommen U, Schäfers C. GLP-code TAM-001/4-50 (2008) Zooplankton community level study with a fungicide product in semi-realistic aquatic microcosms. 136 pages (analytical report separately).

82) Hommen U, Schäfers C, Ebke KP, Böhmer W. GLP-code GAB-016/4-52 (2008) Aquatic mesocosm study on the effects of an insecticide product in an enclosure system in Homberg/Ohm. 116 pages + appendices A-N.

Extended fish studies in microcosms

83) Schäfers C. GLP-code AGR-001/4-51 (1999) Prolonged study on the effects of an respiratory chain inhibitor (acaricide, insecticide) on fish and aquatic invertebrates under indoor semi-realistic conditions in aquatic microcosms. 65 pages + appendices A-E (119 pages).

84) Wellmann P, Schäfers C. GLP-code BAT-001/4-51 (2004) Prolonged fish study on the effects of an insecticide in indoor semi-realistic microcosms. 73 pages + analytical part + appendices + CD.

85) Wellmann P, Schäfers C. GLP-code FEI-015/4-51 (2006) Prolonged fish study on the effects of a herbicide in indoor semi-realistic microcosms. 59 pages.

Fish Full Life Cycle Studies

86) Schäfers C. GLP-code CYA-004/4-61 (2001) Zebra fish (Danio rerio), full life cycle test with sediment, static approach (herbicide, effects following bioaccumulation). 158 pages.

87) Schäfers C. GLP-code BAS-014/4-61 (2003) Zebrafish (Danio rerio), static full life cycle test with sediment (fungicide, effects by endocrine disruption). 239 pages.

88) Schäfers C. GLP-code BAS-015/4-61 (2003) Zebrafish (Danio rerio), static full life cycle test with sediment (insecticide, effects following bioaccumulation). 122 pages.

89) Schäfers C. GLP-code BAS-018/4-61 (2004) Zebrafish (Danio rerio), static full life cycle test with sediment (effects following bioaccumulation). 125 pages.

90) Schäfers C. GLP-code BAY-020/4-62 (2004) Zebra fish (Danio rerio), Two-Generation Test with a pesticide metabolite, flow-through conditions. 74 pages.

91) Schäfers C. GLP-code FEI-016/4-61 (2007) Zebrafish (Danio rerio), static full life cycle test with sediment (fungicide, effects by endocrine disruption). 172 pages.

92) Schäfers C. GLP-code MAK 002/4-61 (2008) Zebrafish (Danio rerio), static full life cycle test with sediment (fungicide, effects by endocrine disruption). 174 pages.

93) Teigeler M. GLP-code ISA 001/4-61 (2008) Zebrafish (Danio rerio), Two-Generation Test with a fungicide (effects by endocrine disruption). 236 pages.

Acknowledgements

I wish to thank the following persons who considerably influenced my scientific background and the direction and motivation of my work.

Prof. Alfred Seitz brought me in contact to population ecology and gave me the opportunity to apply concepts in experimental work. He always was open for all kinds of scientific questions and keen to contribute to solutions.

Prof. Roland Nagel introduced me to aquatic ecotoxicology and the challenges and needs of applied research. Gerhard Görge represented thorough scientific work combined with cordial partnership. Together with him, Uwe Ensenbach, Sabine Zok and Achim Schmitz formed the core of the Mainz University working group of ecotoxicology I had the luck to be part of. We had intensive discussions on scientific and private issues that helped us to develop the profile needed for our tasks.

In my first professional position, Joachim Sonnenberg delegated high responsibility in planning, organizing and performing scientific courses as well as laboratory work in Quedlinburg just after the German reunification. I will never forget the impressive acceptance by and confidence of the participants in the courses in spite of their very different background.

Annette Fliedner and Reinhard Debus offered the opportunity of a post-doc position at the Fraunhofer IUCT. Prof. Werner Klein supported my entry and development as head of laboratory, group and department. He was always straightforward and confident and is still a valuable consultant in all questions related to Fraunhofer work and strategies. Prof. Klein coined my view of the Fraunhofer task as scientific mediator between industry and authorities.

All the presented work was only possible by heading an excellent team of technicians and scientific colleagues with outstanding skills. Especially Uwe Boshof, Josef Greve, Elisabeth Hardebusch, Katja Mock, Pamela Schulte, Annette Fliedner, Udo Hommen, Christian Schlechtriem, Matthias Teigeler and Andrea Wenzel contributed a lot of practical ideas, idealism, time and effort in successful client orientated applied research. We could make use of Udo’s internationally acknowledged expertise in ecological modeling and statistics.

Finally yet importantly my family Petra, Ruth, Simon and Judith gave me the support needed. Warm thanks for your patience and gift of time that we could have spent together.

The presented work was performed by Christoph Schäfers and his staff at the Fraunhofer Institute for Molecular Biology and Applied Ecology IME in Schmallenberg, Germany. This summarizing report serves as a habilitation thesis in ecotoxicology at the University Koblenz-Landau.

We developed and improved experimental approaches to identify the ecological effects of chemical exposure, aiming to minimize uncertainties of regulatory assessments. For hazard assessment of substances, projects on invertebrate species sensitivity distributions in ground and surface water, (endocrine) effects on sensitive fish life stages and performances, and effects on aquatic communities in large-scale microcosms are presented. For water quality evaluation, habitat-specific approaches and multivariate methods to aquatic invertebrate community monitoring are shown. General conclusions are followed by an outlook to current and future research issues, especially with respect to a landscape-level risk assessment.

ISBN 978-3-8396-0542-4

9 783839 605424