Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

02-May-2006

English - Or. English

ENV/JM/MONO(2006)17
Unclassified

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

OECD SERIES ON TESTING AND ASSESSMENT
Number 53

GUIDANCE DOCUMENT ON SIMULATED FRESHWATER LENTIC FIELD TESTS (OUTDOOR MICROCOSMS AND MESOCOSMS)

Ms. Laurence MUSSET
Tel: +33 (0)1 45 24 16 76; Fax: +33 (0)1 45 24 16 75; Email: laurence.musset@oecd.org

JT03208281

English - Or. English

OECD Environment Health and Safety Publications

Series on Testing and Assessment

No. 53

GUIDANCE DOCUMENT ON SIMULATED FRESHWATER LENTIC FIELD TESTS (OUTDOOR MICROCOSMS AND MESOCOSMS)

Environment Directorate
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris, April 2006

Also published in the Series on Testing and Assessment:

No. 1, *Guidance Document for the Development of OECD Guidelines for Testing of Chemicals (1993; reformatted 1995)*

No. 2, *Detailed Review Paper on Biodegradability Testing (1995)*

No. 3, *Guidance Document for Aquatic Effects Assessment (1995)*

No. 4, *Report of the OECD Workshop on Environmental Hazard/Risk Assessment (1995)*

No. 5, *Report of the SETAC/OECD Workshop on Avian Toxicity Testing (1996)*

No. 6, *Report of the Final Ring-test of the Daphnia magna Reproduction Test (1997)*

No. 7, *Guidance Document on Direct Phototransformation of Chemicals in Water (1997)*

No. 8, *Report of the OECD Workshop on Sharing Information about New Industrial Chemicals Assessment (1997)*

No. 9, *Guidance Document for the Conduct of Studies of Occupational Exposure to Pesticides during Agricultural Application (1997)*

No. 10, *Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data (1998)*

No. 11, *Detailed Review Paper on Aquatic Testing Methods for Pesticides and industrial Chemicals (1998)*

No. 12, *Detailed Review Document on Classification Systems for Germ Cell Mutagenicity in OECD Member Countries (1998)*

No. 13, *Detailed Review Document on Classification Systems for Sensitising Substances in OECD Member Countries 1998)*

No. 14, *Detailed Review Document on Classification Systems for Eye Irritation/Corrosion in OECD Member Countries (1998)*

No. 15, *Detailed Review Document on Classification Systems for Reproductive Toxicity in OECD Member Countries (1998)*

No. 16, *Detailed Review Document on Classification Systems for Skin Irritation/Corrosion in OECD Member Countries (1998)*

No. 17, *Environmental Exposure Assessment Strategies for Existing Industrial Chemicals in OECD Member Countries (1999)*

No. 18, *Report of the OECD Workshop on Improving the Use of Monitoring Data in the Exposure Assessment of Industrial Chemicals (2000)*

No. 19, *Guidance Document on the Recognition, Assessment and Use of Clinical Signs as Humane Endpoints for Experimental Animals used in Safety Evaluation (1999)*

No. 20, *Revised Draft Guidance Document for Neurotoxicity Testing (2004)*

No. 21, *Detailed Review Paper: Appraisal of Test Methods for Sex Hormone Disrupting Chemicals (2000)*

No. 22, *Guidance Document for the Performance of Out-door Monolith Lysimeter Studies (2000)*

No. 23, *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures (2000)*

No. 24, *Guidance Document on Acute Oral Toxicity Testing (2001)*

No. 25, *Detailed Review Document on Hazard Classification Systems for Specifics Target Organ Systemic Toxicity Repeated Exposure in OECD Member Countries (2001)*

No. 26, *Revised Analysis of Responses Received from Member Countries to the Questionnaire on Regulatory Acute Toxicity Data Needs (2001)*

No 27, *Guidance Document on the Use of the Harmonised System for the Classification of Chemicals Which are Hazardous for the Aquatic Environment (2001)*

No 28, *Guidance Document for the Conduct of Skin Absorption Studies (2004)*

No 29, *Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media (2001)*

No 30, *Detailed Review Document on Hazard Classification Systems for Mixtures (2001)*

No 31, *Detailed Review Paper on Non-Genotoxic Carcinogens Detection: The Performance of In-Vitro Cell Transformation Assays (draft)*

No. 32, *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies (2000)*

No. 33, *Harmonised Integrated Classification System for Human Health and Environmental Hazards of Chemical Substances and Mixtures (2001)*

No. 34, *Guidance Document on the Development, Validation and Regulatory Acceptance of New and Updated Internationally Acceptable Test Methods in Hazard Assessment (2005)*

No. 35, *Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies (2002)*

No. 36, *Report of the OECD/UNEP Workshop on the use of Multimedia Models for estimating overall Environmental Persistence and long range Transport in the context of PBTS/POPS Assessment (2002)*

No. 37, *Detailed Review Document on Classification Systems for Substances Which Pose an Aspiration Hazard (2002)*

No. 38, *Detailed Background Review of the Uterotrophic Assay Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Assay (2003)*

No. 39, *Guidance Document on Acute Inhalation Toxicity Testing (in preparation)*

No. 40, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures Which Cause Respiratory Tract Irritation and Corrosion (2003)*

No. 41, *Detailed Review Document on Classification in OECD Member Countries of Substances and Mixtures which in Contact with Water Release Toxic Gases (2003)*

No. 42, *Guidance Document on Reporting Summary Information on Environmental, Occupational and Consumer Exposure (2003)*

No. 43, *Draft Guidance Document on Reproductive Toxicity Testing and Assessment (in preparation)*

No. 44, *Description of Selected Key Generic Terms Used in Chemical Hazard/Risk Assessment (2003)*

No. 45, *Guidance Document on the Use of Multimedia Models for Estimating Overall Environmental Persistence and Long-range Transport (2004)*

No. 46, *Detailed Review Paper on Amphibian Metamorphosis Assay for the Detection of Thyroid Active Substances (2004)*

No. 47, *Detailed Review Paper on Fish Screening Assays for the Detection of Endocrine Active Substances (2004)*

No. 48, *New Chemical Assessment Comparisons and Implications for Work Sharing (2004)*

No. 49, *Report from the Expert Group on (Quantitative) Structure-Activity Relationships [(Q)Sars] on the Principles for the Validation of (Q)Sars (2004)*

No. 50, *Report of the OECD/IPCS Workshop on Toxicogenomics (2005)*

No. 51, *Approaches to Exposure Assessment in OECD Member Countries: Report from the Policy Dialogue on Exposure Assessment in June 2005 (2006)*

No. 52, *Comparison of emission estimation methods used in Pollutant Release and Transfer Registers (PRTRs) and Emission Scenario Documents (ESDs): Case study of pulp and paper and textile sectors (2006)*

No. 53, *Guidance Document on Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms) (2006)*

No. 54, *Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application (2006)*

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 30 industrialised countries in North America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in ten different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides and Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and the Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (http://www.oecd.org/ehs/).

This publication was produced within the framework of the Inter-Organisation Programme for the Sound Management of Chemicals (IOMC).

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The participating organisations are FAO, ILO, OECD, UNEP, UNIDO, UNITAR and WHO. The World Bank and UNDP are observers. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

FOREWORD

The history of the Guidance Document on Simulated Freshwater Lentic Field Tests dates back to the 4th Meeting of the Working Group of the National Coordinators of the Test Guidelines Programme (WNT), held in 1993, where National Co-ordinators agreed to initiate the project. The first draft was developed by the Aquatic Model Ecosystem Advisory Committee, under the auspices of the Society of Environmental Toxicology and Chemistry (SETAC); it was based on the scientific consensus made at a series of workshops convened by SETAC in 1991 and 1992. This draft was circulated to the WNT for comments in July 1996.

In order to address the comments made, the draft was reviewed and two international workshops were held in 1999 and 2000. In June 2000, a revised draft was circulated to a group of experts, identified at the international workshop in 2000. Comments from these experts were incorporated into the text. The revised draft was submitted to the 15th WNT Meeting in May 2003. The WNT felt that general discussion was still needed before finalising the technical details of the Guidance Document, and agreed on the following two-step process:

• First circulate the slightly revised version of the Guidance Document to the National Coordinators for their comments on the scope and the approaches of the document;
• Then revise the draft and circulate again for detailed comments.

The circulation for comments on the scope and approaches took place in August 2003. Comments from member countries were received and a teleconference of experts was held in January 2004. The experts agreed that the approaches of the Guidance Document to focus on aquatic effects with some reference to fate properties are appropriate, and provided advice to the Secretariat on the revision of the draft document taking into account the comments from the member countries.

The Secretariat circulated a revised draft in July 2004 for more detailed comments and revised the document subsequently. The 17th WNT Meeting approved the draft Guidance Document in April 2005, subject to the addition of references cited in a paragraph. This issue was resolved by written procedure in September 2005.

This document is published on the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology.

TABLE OF CONTENTS

INTRODUCTION

1. Outdoor microcosms and mesocosms are useful in risk assessment when lower-tier and higher-tier laboratory studies (single-species and multi-species) indicate potential risks. They can be an important tool in bridging the gap between these studies and attempts to understand, predict and confirm what may occur in the natural environment subject to these risks. Every microcosm or mesocosm study should be designed to test a specific hypothesis using information gained in previous steps of the risk assessment. This makes every microcosm or mesocosm study unique in at least some aspects of its design. Guidance for conducting microcosm or mesocosm studies is therefore necessarily generic and flexible (Campbell *et al.* 1999).

2. Although this guidance document is mainly based on experiences with testing of plant protection products (PPP), in principle, it applies to other groups of toxic substances as well (e.g., wastewater, industrial chemicals). However it should be noted that suitable exposure design for substances other than PPPs can differ (e.g., steady contamination instead of single application)."

3. This document was originally based on consensus methods proposed by experts at a series of meetings convened within Europe and North America (SETAC-Europe, 1991; SETAC/RESOLVE, 1991; EWOFFT, 1992; Hill, *et al.*, 1994). Also, recent developments and experience from international workshops (Campbell *et al*. 1999; Giddings *et al.* 2002) have been incorporated into this document.

4. One important reason to perform a microcosm or mesocosm study is to obtain more knowledge about the ecological relevance of effects identified in laboratory studies. The studies can therefore include a variety of species, functional groups or habitat types. Interpretation of these studies focuses on effects at the community and ecosystem level, potential indirect effects and the recovery potential of sensitive organisms. A second important reason for conducting a microcosm or mesocosm study is to measure effects of the chemical under more environmentally realistic exposure conditions. Such conditions could include the influence of partitioning to sediments and plants, photolysis, and other processes that may influence the fate of the chemical. Moreover, microcosm or mesocosm studies incorporate natural abiotic conditions (temperature, light, pH, *etc.*) that may influence the response of certain organisms (Campbell *et al.* 1999). Another reason for performing a microcosm or mesocosm study is that by conducting this test one can assess and study the impact on the structural and functional attributes of natural ecosystems.

5. Although the scope of this guidance document mainly relates to determining environmental effects, these cannot be separated from fate properties. In this regard, concentrations are measured at appropriate time intervals in the compartments of concern. Studies purely aiming at fate properties are not specifically addressed in this guidance document, but have been dealt with elsewhere (SETAC/RESOLVE, 1991; SETAC, 1995; Campbell et al. 1999).

GENERAL DESCRIPTION OF EXPERIMENTAL SYSTEMS

6. Outdoor mesocosm or microcosm studies can be performed with artificial tanks or ponds or by enclosing parts of existing ecosystems. When constructing a mesocosm facility it is a good idea to include one or more 'supply ponds' which can be used as a source of water, sediment and organisms. A size of 1 to 20 m^3 is usually regarded as appropriate for outdoor meso- or microcosm studies. Where planktonic species are the main concern, microcosms of between approximately 100 and 1000 litres can readily be used. It is also possible to use much larger mesocosms, although they may be more logistically demanding than smaller systems (eg Kersting & van Wijngaarden, 1999). The size to be selected for a meso- or microcosm study will depend on the objectives of the study and the type of ecosystem that is to be simulated. In general, studies in smaller systems (about 1 - 5 m^3) are more suitable for shorter studies of up to three to six months and for studies with smaller organisms (e.g. planktonic species). Larger systems are more appropriate for longer studies (e.g. 6 months or longer). The average depth of test systems depends on the objectives of the study but, generally, should be between 30 cm and 1 m. In small systems, large daily temperature fluctuations in the test systems can be buffered by partial burial in the ground or immersion in a pond.

7. It is difficult to distinguish between microcosms and mesocosms based on spatial scale (size dimensions or volume) as there is some overlap and differences of opinion. However, Cooper and Barmuta (1993) makes this distinction, spatially and temporally, to a certain extent: microcosms are 10^{-3} to 10 m^3 in size, while mesocosms are 1 to 10^4 m^3. Larger mesocosms may be considered to be equivalent to whole or natural systems (10^3 to 10^8 m^3 in size). These cut-off values along with other information are included in Table 1.

8. Guidance in this document is applicable and relevant only to lentic (static) microcosm or mesocosm systems and not to lotic (flowing) systems, such as constructed artificial model streams, actual streams of various sizes found in the environment or segregated or enclosed portions of actual streams. Furthermore, guidance in this document is restricted to outdoor systems. Indoor, laboratory, or "bench top" microcosms are not considered. This guidance document is not applicable to most mesocosms that are large enclosed portions of large natural ecosystems e.g. vertical portions of the water column in a large lake enclosed by large-sized bags or enclosures. Instead, this guidance document is applicable to littoral enclosures of smaller natural ecosystems (e.g., ponds and small lakes), in situations where an area is subdivided into two or more smaller areas, or where smaller-sized bags or enclosures are used.

INITIAL CONSIDERATIONS

9. Before any microcosm or mesocosm test is conducted, clear objectives should be defined in order to determine which relevant endpoints are and which experimental design (e.g. level of replication, number of treatments) is appropriate. It is the responsibility of the study director to demonstrate that the system is appropriate for achieving the objectives of the study. It may be useful to discuss and agree on the protocol with the relevant authorities evaluating the test results. Any available information should be carefully reviewed, and preliminary laboratory testing should be undertaken when essential information for test design is missing. Factors to be considered include:

i) Effects properties: the core ecotoxicological data that are always required for registration and other higher tier studies (e.g. additional single species, population-level studies, indoor multi-species studies) can be used to determine the primary concerns to be investigated. For example, data on the sensitivity of aquatic species can help to focus those populations and communities which should be studied in more detail.

ii) It should be determined what the derived parameters of interest are for each endpoint (e.g. EC_{50}, EC_x or NOEC).

iii) The level of precision that is to be obtained for derived estimates, or the desired power of a relevant hypothesis should be determined as part of determining the objective of the study.

iv) The size of effects which is considered of ecological significance should be determined, relative to the endpoints of concern and the characteristics of the test species such as generation time, reproductive and migratory endpoints.

v) Depending on the objectives following on from point (a) to (d), it will be possible to determine:

- Number of treatments and choice of doses
- How treatments are to be assigned to the mesocosms (random or block design, etc.)
- Number of replicated mesocosms per treatment
- Organisms to be sampled, the size of the sample and how sampling should be carried out.

vi) The method of statistical analysis should be determined as part of the setting of the objectives. (Chapman & Maund, 1996).

vii) An appropriate exposure regime should be established in order to meet the objectives of the study. Questions that should be addressed include: (i) what are the expected routes of entry of the chemical into aquatic systems (e.g. spray drift or surface run-off); (ii) what is the frequency and timing of the entry of the chemical.

viii) Physical-chemical and fate properties (for example, solubility, vapour pressure, octanol/water partition coefficient, adsorption coefficients, hydrolysis and photolysis rates, and biodegradability) should be previously ascertained alongside biological information in order to select sampling times and identify ecological components at greatest risk. A valid analytical method should be available before performing the microcosm or mesocosm test.

ix) If it is felt necessary to examine the relationship between loading and resulting exposure concentrations in various compartments or strata in addition to lower-tier fate data, then chemical input and fate modelling or field studies specifically designed to determine the fate and behaviour of the chemical can be used. There are a large number of models that can be used to derive initial estimates of persistence and concentration-time functions (See FOCUS and ECFRAM reports for an overview of models). Appropriate fate studies may be necessary before starting the microcosm or mesocosm test.

x) Information on the use patterns of the product is necessary.

xi) Information on the likely major metabolites should be considered.

PRINCIPLES OF THE TEST

10. Microcosms and mesocosms are constructed to simulate parts of natural aquatic ecosystems. These are generally established either by collecting organisms and placing them in tanks or artificial ponds or by enclosing parts of existing ecosystems. The test system thus contains a 'naturally' developed aquatic community, usually containing naturalised sediment and appropriate organisms such as zoo- and phytoplankton, pelagic and benthic macro-invertebrates and macrophytes. It may be appropriate to add certain organisms (eg. fish) from external sources. It should be noted that when parts of larger ecosystems are enclosed, the more limited space and boundaries may influence and affect the organisms and processes occurring in the microcosms or mesocosms, thus causing divergence from the larger ecosystem.

11. Environmental risk assessment of chemicals is usually aimed at protection of communities (including the populations of each contributing species). Thus, it is desirable for the test system to be broadly representative. In practice, however, any test system can only simulate parts of one whole ecosystem, and extrapolations have to be made when transferring the results of a microcosm/mesocosm study to other types of ecosystems or ecosystem components that have not been tested (e.g., recovery of semi- or univoltine species). In principle, the relevance of the test system for the protection purposes in question has to be discussed in the context of data interpretation.

12. Application of the test substance can be made by direct addition of the chemical to water or by simulating the actual route of exposure (e.g. by application of spray, to simulate spray drift or by application of a soil-water slurry to simulate erosion runoff). In general it is recommended to have an exposure regime that will allow determination of a concentration-response relationship if feasible. Exposure is expressed in terms of the concentrations of the chemical in water (toxicological approach), as in a laboratory toxicity test. This approach is preferred over a simulation approach where exposure is expressed in terms of the amount of the chemical added per unit surface area or volume (Giddings *et al*. 2002). The toxicological approach facilitates the use of effect data in different exposure situations (frequently, exposure situations emerge which were not initially anticipated), although the simulation approach may be useful to confirm the acceptability of a particular use pattern. (Also see paragraph 41.)

13. Microcosm or mesocosm studies should preferably be designed so that a concentration-response relationship may be identified over a range of ecotoxicologically relevant concentrations encompassing those concentrations that reflect exposure in the field (cf. para 41).

14. Structural endpoints and functional endpoints should be considered. Structural endpoints relate to the abundance and biomass of all populations and their spatial, taxonomic and trophic organisation (Brock and Budde, 1994). Functional endpoints are related to all aspects of non-living materials processed by the structure – i.e. nutrient levels, oxygen levels, respiration rate, mineral concentrations, pH, alkalinity, conductivity and organic material content (Kersting 1994; Brock et al. 2000a). In studies performed with substances that are not herbicides, the functional aspects could also be documented as conditions of the study rather than as endpoints of the study. The microcosm or mesocosm study should focus on taxonomic groups that lower-tier risk assessments have identified as being of concern, as structurally or functionally important, or as exhibiting sensitivity to the test substance.

15. Determining rate and extent of recovery of affected taxa can be crucial in the design of microcosm or mesocosm studies. Looking at recovery is one of the key differences between microcosm and mesocosm studies and other higher tier studies and requires substantial ecological knowledge to interpret. If looking at recovery is an objective of the study, the experimental design should be such that recovery

can be observed. A sufficiently long post-treatment period of two to three generation times has to be foreseen to allow the detection of repopulation. (Giddings *et al*. 2002)

16. The data handling and statistical methods that are to be used to analyse the data should be built into the design of the study.

VALIDITY OF THE TEST

17. Because of the complexity and variability of the test systems, this guidance document does not provide absolute validity criteria. However, the validity of a study can be evaluated in light of the following conditions.

- The study should focus on endpoints of organisms that are potentially at risk, which should have been identified in other lower- or higher-tier studies.
- Ideally, when concentration-response is studied, a clear effect level for at least the organisms of concern should be included and at least one level that causes no effects that are considered ecologically significant. (based on expert consideration of ecological function and recovery).
- Variability should be as small as possible. If the variation between replicates is high, then the conclusions drawn from the study are less robust.
- The amount of test material applied and the exposure concentration in the water column should be determined analytically at t=0 (start of exposure). Whether or not concentrations are measured in other matrices, depends on the objectives of the study.
- The duration of the study should be appropriate to the life-cycle of the organisms of interest and the time needed for their recovery, if that is an objective of the study.

DESCRIPTION OF THE METHOD

Experimental systems

18. The microcosm or mesocosms can be constructed from any natural substrate or inert material, such as concrete (sealed appropriately), fibre- or plexiglass or stainless steel. Systems can also be lined with inert plastics to prevent exchange of water with the surroundings. Care should be taken to prevent leaching of plasticisers into the test waters; an epoxy paint may be necessary. In some locations, it may be advisable to cover the test systems with a net to prevent disturbance of the test systems by large birds such as herons or ducks.

Re-use of ponds

19. Re-use of ponds after treatment with toxic chemicals depends on chemical characteristics, particularly the persistence of the chemical, and on biological variability. For non-persistent chemicals there may be no problem if it can be demonstrated that there are no longer any toxic residues present in the water column and in the sediment and the systems can all be readily returned to closely comparable biological systems. Alternatively, the ponds can be drained and left empty for a period of time, or scraped out and relined with new sediments (Hill *et al*, 1994).

Sediment

20. Sediments should always be included in the test systems because they provide an important buffering element to the systems. Sediment can be collected from supply ponds or natural systems, in which case care should be taken to ensure that organisms which could interfere with the objectives of the study (e.g. fish or highly invasive or exotic macrophytes) are not accidentally introduced. Alternatively, soil can be used as the test system sediment, provided that it has been sufficiently conditioned to have aquatic sediment-like properties. Generally this requires a maturation period of immersion of 3 months or more and addition of a small inoculum of aquatic sediment to encourage the development of suitable microflora. Sediment collected from an uncontaminated reference site will contain indigenous flora and fauna which can be used to establish pond communities. If soil is used, then additional organisms may need to be added to develop suitable communities for the study. The sediment should be characterised by analysis of chemical residues, including metals and by determination of the particle size distribution, organic matter content. N and P content, cation exchange capacity, pH and organic carbon content could be quantified to further characterise the sediment. Prior to adding sediment to the individual units, it should be thoroughly mixed to ensure an even distribution of material and benthic organisms and thereby promote inter-replicate similarity. The sediment is added to the microcosm or mesocosms prior to adding water. The depth of sediment in each test system should be > 5 cm.

Water

21. If possible, at least part of the water, and preferably all of it, should originate from the zone where the sediment and its organisms were collected. As with the sediment, water used in the microcosm or mesocosm should be characterized for chemical contaminants, nutrients, pH, hardness, dissolved oxygen, and turbidity. When adding water to the test systems, it is advisable to do so gently so that the sediment quickly settles. If large systems are used, or the study director suspects that there may be high levels of variability between the replicates, it is possible to interchange water between microcosm or mesocosms. However, mixing should be discontinued sufficiently before treatment to ensure that the test systems are relatively stable without water exchange between the systems. After dosing, no water exchange should occur among the microcosm or mesocosms.

22. Water level should be maintained at similar levels (within about 20% of the original level) for the duration of the study. This should be done by replacing evaporated water with a well source, ‘conditioned’ mains water, or filtered pond water (preferably water with similar properties to that used to establish the test systems). In very wet weather conditions, test systems can be covered to prevent overflow. In emergencies, water can also be removed, but if this is after treatment, an estimate of the amount of residues removed should be obtained. However, the addition of small volumes of unfiltered water, containing plankton, may simulate natural immigration and water exchange, and therefore may be done if the study objective is to show recolonisation rather than intrinsic recovery resulting from the reproduction of surviving organisms.

Organisms to include

23. The microcosm or mesocosm study should focus on taxonomic groups that lower-tier risk assessments have identified as being of concern. Besides naturally colonised sediment, the test system is usually a naturally developed aquatic community with appropriate organisms such as zoo- and phytoplankton, periphyton, bacteria, macrophytes, pelagic and benthic macro-invertebrates. To develop communities suitable to meet the study objectives, it may also be acceptable to add organisms from appropriate external sources. In most cases, it is recommended that macrophytes should be included in the test systems, even if the objective is only to study phyto- and zooplankton. This is because macrophytes are an important structural and functional part of aquatic ecosystems, providing habitat for organisms and contributing to macro- and micronutrient cycling and influencing physico-chemical conditions. With regard to the inclusion of fish, see paragraph 27.

Macrophytes

24. In microcosm or mesocosm studies where macroinvertebrates are the endpoints of major interest, macrophytes should be present, since they provide refugia for the microcosm or mesocosm fauna. Their presence is likely to encourage greater system stability and greater diversity of algae and invertebrates. For auxin-stimulating herbicides it is essential to include relevant macrophyte species, preferably those that develop roots (Brock *et al.* 2000). Another reason to include macrophytes in these studies is that effects on macrophytes can have indirect effects on the animal communities because of the interdependencies described above.

25. Natural colonisation is acceptable; propagules and seeds will be found within most introduced sediments. Planting macrophytes is also a possibility and this also has benefits because it promotes more uniformity in the species structure, abundance and distribution of the fauna and flora. Planting mature macrophytes (e.g. from a supply pond) also increases the complexity of the system by introducing new micro-habitats and enhances the rate at which the system can 'mature'. Macrophyte development should be managed to ensure that the objectives of the study are met. For example, some floating (e.g. *Azolla* or *Lemna*) or submerged (e.g., *Elodea*) species can overdominate a system and significantly reduce animal diversity. If the objectives of the study are focused on the plankton, then it may be advisable to maintain areas of open water, limiting macrophyte development to no more than 50% of the bottom area; typically, 25 to 30%. If the study is to focus on macroinvertebrates it may be appropriate to promote the abundance and diversity of submerged macrophytes in order to enhance abundance and diversity of macroinvertebrate species (these are closely ecologically linked). It may also be wise to include emergent species, if aquatic insects are of concern because many aquatic insect species use these plants for emergence or egg-laying.

Invertebrates

26. Benthic and planktonic invertebrates will often be added to the microcosm or mesocosms with the sediments and water. Invertebrates typically studied in detail include the zooplanktonic phyla Rotifera and Arthropoda (Branchiopoda: Cladocera, Copepoda), and zoobenthic organisms from the phyla Annelida (Oligochaeta, Hirudinea), Mollusca (Gastropoda, Bivalvia), Arthropoda (Insecta e.g. Coleoptera, Diptera, Ephemeroptera, Hemiptera, Odonata, Trichoptera; Crustacea e.g. Isopoda, Amphipoda, Ostracoda, Decapoda), and Platyhelminthes (Turbelleria). Also, epibenthic invertebrates, and those invertebrates found growing on macrophytes (e.g. Bryozoa) could be studied in microcosm or mesocosm tests. Should it be necessary, any group can be added to the microcosm or mesocosms from field collections or laboratory cultures at the start of the study. Care should be taken to ensure a homogeneous distribution between microcosm or mesocosms prior to dosing. This can be accomplished by compositing samples and by using sample splitters.

Fish

27. Inclusion of free-living fish is not generally recommended in smaller test systems, particularly where effects on zooplankton and macro-invertebrates are key study end-points. If direct effects on fish are of concern, laboratory single species studies are usually more appropriate than microcosm or mesocosm studies. The reason is that, in a confined ecosystem, fish tend to have an unnaturally large influence on the rest of the system. In tests of long duration or high fish biomass density, predation by fish may influence invertebrate responses to the chemical. However, fish are appropriate where the indirect effect of a chemical on fish production is of interest. For example, when the toxicity of the chemical is expected to have substantial impact to one or more invertebrate groups and thereby reduce the available food supply of the fish population."

28. If one objective of the study is to look at effects on fish population, then the larger mesocosm enclosures are recommended. It may be useful to stock the microcosm or mesocosm with a low density of adults and remove adults and larvae after spawning. However, the life stage, number and biomass of fish added also depend on the purpose of the test. For example, should the emphasis be on an insecticide, larval fish may be added to monitor their growth in relation to the invertebrate food base. If applicable, the fish population should have a natural demographic structure and should not exceed the "carrying capacity" of the test system (Brock & Budde, 1994). Biomass densities should generally not exceed 2 g per m^3 (Touart, 1988; Fairchild *et al.,* 1992).

29. If fish are included in the test system, they should be added after the test system is reasonably stabilised. Typically, this is from one to four weeks. When free-living fish are included and only direct effects are of concern, it may also be possible to introduce caged fish. If free-living fish are included, it may be advisable to include an invertebrate 'refuge' where fish are excluded in order to maintain reasonable numbers of unconsumed invertebrates.

30. Fish used should preferably be representative for fish species living in the habitat or type of ecosystem which is of concern for the investigation. In order to avoid contamination of native fish habitats, only endemic fish species should be used in mesocosm tests. Examples of appropriate fish species include bluegill sunfish (*Lepomis macrochirus*) fathead minnow (*Pimephales promelas*), sticklebacks (*Gasterosteus aculeatus*), mosquito fish (*Gambusia affinis*), carp (*Cyprinus carpio*), golden orf (*Leuciscus idus)* or rainbow trout (*Oncorrynchus mykiss)*. Which fish species is most appropriate depends on the objective of the test and the size of the test system.

Maturation time

31. The period of time the microcosm or mesocosms are adapted prior to chemical dosing should ensure the development of a community of populations representative of field populations in terms of age- and sex-structure. This period will vary with the size of the system and the origin of the introduced sediment/water. A sufficient degree of diversity of the system for the purpose of the study and a certain homogeneity between replicates should have been achieved before the study can begin.

TEST DESIGN

32. Appropriate test design depends on the purpose of the test: establishment of dose (concentration) response relationship, comparison with laboratory-derived NOEC, etc.

33. An exposure-response experimental design with replication allows wider use of the data under different conditions and for different regulatory requirements. In this design, in general a microcosm or mesocosm study should include at least three, and preferably five concentrations, with at least two replicates per concentration. More concentrations may be required, depending on the slope of the dose-response curve, for taxa of interest. Statistical techniques may be stronger and the power to detect differences is increased with more replicates, since replication reduces uncertainty in interpretation of results, and because test system variability can be better accounted for. However it is also possible to design valid studies with a large number of unreplicated treatment levels. In such cases, in order to determine a statistically robust EC_x intensive range finding tests should be conducted to determine treatment levels where responses for the critical endpoints in the range of 10 to 90% effects occur. The decision either to favour more replicates of each concentration (to calculate NOEC) or to prefer a concentration-response test design with less replicates and an increased number of concentration levels depends on the scientific questions to be answered and on the specific characteristics of the established ponds (Chapman & Maund, 1996; Campbell *et al.* 1999; Giddings *et al*. 2002).

34. In designing the microcosm or mesocosm study it may be helpful to consult a statistician to help determine which test design is required if an effect in a particular set of measured parameters is to be determined with a specific power. This will be a function of the replicate number and variability of the measurements. A design optimal for one variable will not necessarily be appropriate for another. The importance of focussing on critical endpoints cannot be overemphasized (Chapman & Maund, 1996).

35. The selected concentrations should generally be based on those expected to cause effects. If possible, this should include the maximum predicted environmental concentration (PEC). Where feasible and relevant, the complications caused by multiple applications should be avoided. The selection of treatment levels should aim to include at least one concentration that will cause no ecologically significant effects and at least one that will cause clear effects. These concentrations can be derived from lower-tier studies.

36. Data about the natural variability of the biological assemblage of the microcosm or mesocosm system should be available to design the study in such a way that a high number of species can be included.

37. Whether treatments should be assigned to experimental units at random or whether a constrained randomization should be employed, such as the arranging of treatments in replicate blocks, depends on the objective of the study (Chapman & Maund, 1996). As in any experimental treatment using chemicals, care should be taken not to contaminate other microcosm or mesocosms with test material.

Recovery

38. Determining the rate and extent of recovery of affected taxa can be an important factor in the design of microcosm or mesocosm studies. When considering recovery, it is important to understand the potential influence of life-history and dispersal mechanisms of the organisms involved and possible interactions of these with the exposure regime and tests system. For example, if application is carried out during or after the main reproductive period of univoltine organisms (e.g. certain mayfly species), if affected organisms do not have non-aquatic dispersal mechanisms (e.g. amphipod and isopod crustaceans), or if normal seasonal variations (e.g. phytoplankton) mean that the affected organism becomes absent from

controls and treatments, then it may not be possible to demonstrate recovery potential without further experimental manipulation. Furthermore, in order to evaluate recovery, it is important to consider functional parameters (e.g. production) and possible adaptations/increased tolerance in the organisms/communities in response to the stress. In such cases, certain experimental techniques (e.g. bioassays) may be useful to assist in the evaluation of recovery potential.

39. Periodic reintroduction of organisms, eggs or resting stages into model ecosystems can be used to simulate immigration and reproduction. However this process may have practical constraints placed on it (e.g. for seasonal reasons) and may also disturb the resident populations. It is also very important to consider whether such immigration would occur naturally for the species concerned. Alternatively, organisms may be placed in cages within the model ecosystem from time to time, to determine when the conditions in the model ecosystem have become suitable to support survival, growth, and reproduction. Another possibility to demonstrate recovery potential is to collect water or sediment from the test system and bioassay it in the laboratory (Campbell *et al* 1999).

PROCEDURES

Application of the test substance

40. Test substance is added into the test system after the microcosm or mesocosms have stabilized. For pesticides, whether or not to use the active ingredient or the formulation should be determined by earlier laboratory studies. For a generic risk assessment, use of the active ingredient is preferred, unless it is difficult to work with or if the formulation is significantly more toxic. Formulations should generally be used in studies simulating spray drift, but are inappropriate for studies simulating surface run-off. Regulatory objectives will help determine whether the active ingredient or a formulation should be used (Giddings *et al.* 2002).

41. Two approaches to applying the test substance to the microcosm or mesocosm can be identified.

i) The first resembles a 'toxicological approach' which consists of direct application of the test substance to the water, usually including mixing, to achieve a uniform distribution. Exposure is then expressed in terms of the concentration of the test substance in the water as in a standard toxicity test. However whether this test approach is appropriate depends of the chemicals fate and partitioning properties of the test substance. If this testing approach is used then it is possible to establish ECx- or NOEC-values.

ii) The second approach resembles a 'simulation approach', mimicking the route of entry of the test substance. For pesticides, this amounts to simulating its entry into the water body as under good agricultural practice. For example: drift and direct overspray can be simulated by doing a spray application or erosive run-off can be simulated by making a soil-water slurry. In this case exposure is expressed in terms of the amount of pesticide added per unit surface area or volume (the loading rate). If this testing approach is used, then it may be possible establish a loading giving raise to no effects or a certain effect level.

In both cases concentrations of the test substance established in the water should be analytically determined. The loading (amount of test substance added), the frequency of dosing, and the number of

replicates per treatment necessarily stem from the nature of the chemical, use patterns, routes of entry, variability of the endpoint of concern and objectives of the study.

42. Accounting for multiple exposure events when applying the test substance presents a number of difficulties. Multiple application could be done in the test, however this makes it more difficult to characterise the exposure and interpret the exposure-response relationships.

43. When to apply the test substance depends on how realistic the exposure scenario is to be. It can be argued that it is worst-case to apply the test substance in the spring to midsummer, because the system is assumed to be most sensitive at that time. Applying in the spring also allows a longer time to determine the potential for recovery. Also, if the systems can be assumed to be more sensitive in the spring, the results can be more readily extrapolated to a situation where application is done in the autumn than the other way around. However, there may be a reason to expose the test system in the autumn if it is envisaged that a potential for recovery seen in a spring application has no relevance to the use of the product in the autumn.

Poorly soluble materials

44. Water is by far the preferred carrier for the test substance. A range of solvents that are used as carriers in the application of a test substance, can cause effects on the metabolism of microcosms/mesocosms, even at very low concentrations (Brock *et al.* 2000). If a carrier is used, then the concentrations should be equal across all microcosm or mesocosms, including the controls. Controls with and without a carrier can help to determine whether or not the solvent influenced the test systems.

Sampling

45. Useful information on ecological sampling for population- and community-level effects is available in the literature (Day, 1989; deNoyelles, *et al.*, 1989; Voshell, 1989; Wetzel & Likens, 1991; Fairchild, *et al.,* 1992; Howick, *et al.* 1992). In addition, more specific procedures for lentic microcosm or mesocosm testing are described in numerous papers e.g. by Graney, *et al.,* 1993 and Hill *et al.* 1994. More recent references can be found in Brock *et al*., 2000a and 2000b.

46. In studies where many endpoints are measured more or less simultaneously, it is recommended to assign measurements and samplings to specific locations in the microcosm or mesocosms to avoid mutual influences or disruption of the individual sampling programmes (van *Wijngaarden et al*., 1996).

47. Pre-treatment samples should be taken, for example on day -14, -7 and 0, in order to assess and demonstrate the suitability of the test system. These pre-treatment samples can also be used to perform covariate analyses in order to reduce residual variance among mesocosms (Chapman & Maund, 1996). Sampling continues after treatment for the duration of the test. The total test duration is dependent on the aim of the study, the fate properties of the chemical, recovery times of the populations of concern. Ideally the study should continue long enough to demonstrate the recovery of the affected species.

48. The sampling regime during the exposure period depends on the objectives of the test, the nature of the chemical(s), and the expected distribution of chemical within the microcosm or mesocosm. Table 2 gives an overview of parameters that could be considered in a typical microcosm or mesocosm test. Depending on the type of compound, certain parameters will need to be emphasised more than others. In tests focusing on invertebrate population responses, zooplankton and emergent insects may be sampled weekly. In principle, all samples of the chemical, phytoplankton, zooplankton, and bentic invertebrates should be taken as close together as is practical to strengthen the prediction of the association among these variables.

49. The sampling strategy should ensure that the collection of water samples will not significantly change the microcosm or mesocosm volume. Also the collection of biota specimens should not lower to a significant level the standing stock of the samples species or alter the trophic relationships in the existing foodchains present in the microcosm/mesocosm..

50. If freeliving fish are added to the system, they may be observed at daily or weekly intervals but generally collected upon test termination. There may be occasions in which marked (e.g., electro-tagged) or caged fish are repeatedly sampled for growth measures. However, mortality and effects can increase dramatically for fish repeatedly sampled. Alternatively, in large systems with many fish, it may be possible to take sub-samples to estimate growth rates.

Biological measurement

Phytoplankton and zooplankton

51. The planktonic biota should be sampled with a depth-integrating sampler. In small systems a pump or a plankton net could be used to filter plankton. When using a pump for sampling zooplankton, it should be confirmed that zooplankton (especially the larger members) is not able to avoid the pumps inlet. In the presence of macrophytes, specific techniques are needed to collect zooplankton found in this habitat. Subsamples are used to determine pigment composition, to identify species, and to make cell counts, where appropriate. Population densities are reported as cells (or biomass) per volume. Adult zooplankton are identified to species, where possible and their abundance reported as individuals per volume unit. It should be noted that plankton will be distributed unequally in the water column (e.g., due to daily movements of zooplankton within the water column) and over the surface of the mesocosm. Several zooplankton taxa undergo microhabitat shifts on a diurnal basis; these may be sampled with an integrating sampler. In general, for planktonic samples, depth integrated samples should be used or at least samples need to be taken at the same time and depth in the different replicates. Epi-benthic species may be sampled with a trap designed to collect them (Lozano, *et al.,* 1992).

Periphyton

52. Typically, periphytic biomass and productivity have been estimated using surrogate measures of phytopigments, chiefly chlorophyll-a. Phytopigments (eg chlorophyll-a, phaeophytin) are sampled either from "natural" substrates (e.g., macrophyte surfaces), unglazed ceramic tiles or glass microscope slides placed into racks and subjected to colonisation in the microcosm or mesocosms, for two to four weeks. After exposure, the substrates are scraped and material can be analysed for species composition and abundance, pigment biomass, or ash-free dry-weight. Additional chambers (e.g. light and dark bottles) can be included for estimates of production and respiration, should the emphasis of the test be on factors affecting primary production. The use of colonised substrates allows estimates of colonisation rates and standing crop throughout the test schedule. A series of slides can be pre-exposed within the microcosm or mesocosms and, by collecting slides at frequent intervals, used to monitor changes in periphyton biomass, species composition, and net production. The substrates chosen for use will influence the nature of the periphyton collected. Hence, careful thought should be given to the objectives in estimating primary production or periphytic biomass.

Primary productivity

53. If a chemical is expected to cause direct toxicity to algae (e.g. an herbicide), measures of primary productivity should be used. The most practical method is to measure diurnal oxygen fluctuations. It may also be useful to compare primary productivity with the level of light saturation, in order to explain effects seen. At the same time it is also important to analyse algae populations to a sufficiently low taxonomic

level to be able to detect changes in species composition of primary producers. If possible total system metabolism should be established (Hill *et al*, 1994).

Macrophytes

54. If macrophytes are a study endpoint, e.g. in the case of an herbicide study a system could be set up looking at subsamples for biomass estimates. Care should be taken however to minimise disturbance to the system. It may also be possible to plant macrophyte species of interest in pots placed within the sediment or on the sediment, or suspend them in mid-way the water column. In such cases, the entire pot may be removed for measurement of shoot growth, biomass or photosynthesis. During the study estimates of abundance can be done visually (e.g. by mapping or photographically). Macrophyte growth (e.g., increases in stem length) and biomass should be monitored before senescence starts in late summer.

Heterotrophic component

55. If effects are expected on primary producers such as macrophytes, it may be informative to look at indirect effects on bacteria and detritivores. Specialised techniques are needed to study structural aspects of microbial populations (Wetzel & Likens, 1991). In practice it may often turn out that impact on microcosmorganisms can only be studied by measuring process rates (Giddings *et. al.* 2002).

Macroinvertebrates

56. Macroinvertebrates can be sampled using artifical substrates, nets, direct sampling of sediments, or by using emergence traps. Fast-moving invertebrates can best be sampled directly with a net in the water column. In systems smaller than $10m^3$ it may be advisable not to sample sediments directly. Disturbances to the sediment structure, macrophytes, recycling of nutrients and test material due to sediment sampling may disturb the test system. In these smaller systems, placing sediments in trays prior to application and retrieving the sediment after a period of colonization is one way in which the sediments can be assessed without unduly disturbing the system. To fully characterize the benthic invertebrate fauna, all three methods should be employed in concert. The numbers of collected invertebrates are reported on a per sample basis. Insect emergence rates are reported as numbers emerging per unit of time or per unit area.

Fish

57. Where fish are included, frequent observations should be made at the initiation of the test for dead fish and for abnormal behaviour. Following arrival at the testing facility, fish should be acclimated to the water of the test system for at least a week, before being introduced into the test system. After acclimation, fish should be observed at least weekly and dead fish removed. Fish that die as a result of handling stress or disease may be replaced within the first week of their introduction. At the termination of the test, all fish are collected, counted, measured and weighed. Depending on the test substance and the objectives of the study, fish could be sampled at several time intervals during the study as well, e.g. to measure fish growth. Abnormal growth, external lesions or abnormalities are recorded. Tissue subsamples may be taken for residue analysis, if potential risks of bioaccumulation are indicated in the lower-tier assessments.

58. There is no standard technique presently accepted for determining effects on fish growth and reproduction in microcosm or mesocosms. In designing such a test, it would be necessary to evaluate how best to match biological sampling with the nature of the chemical(s) and also the life history of the organism in question. If the test were to focus on effects on fish early life stage development, the test could be started with eggs or larvae. Test duration would be a function of the fish biomass loading and specific objectives: examples include feeding responses on zooplankton, prey-switching at a critical size, and changes in competitive behaviour as a function of chemical exposure. If the test were to observe

effects on reproduction, adult fish at a low stocking density may be added, allowed to spawn, and subsequently collected with their offspring.

***In situ* bioassays**

59. *In situ* bioassays can provide information on direct effects that can be related to exposure. They may be able to provide information on indirect effects, e.g. in the case of caged fish. *In situ* bioassays can also be used to compare laboratory and microcosm or mesocosm responses in the same species. They should, however, not dominate the system. The effect of the presence of *in situ* bioassays on abundance of free-living populations and the sampling programme should be small (Giddings *et al*. 2002).

Analytical measurements

Water quality

60. Measurement of water quality (e.g. dissolved oxygen, pH, alkalinity, turbidity) and nutrient concentrations must be done to define the ecosystem functioning of the microcosm or mesocosms and to interpret chemical fate and bioavailability.

Analysis of test chemical

61. The study objectives will determine the appropriate sampling and analysis strategy for the test chemical (Giddings, 1994). There are three different reasons for doing chemical analysis: (i) to confirm that the test substance has been accurately applied to the test system (ii) to quantify the chemical exposure and relate it to the ecological responses observed and (iii) to look at the chemical's fate in the aquatic environment under natural or semi-natural conditions.

62. If the test chemical is soluble in water, and is added directly to the water of the microcosm or mesocosm, then measurement of concentrations in water within a few hours after application can provide confirmation of treatment. Vertically integrated water column samples can be taken from a sufficient number of points in the microcosm or mesocosm to allow calculation of the average chemical concentration in the water. A similar approach can be taken to measure insoluble or highly sorbed chemicals that are added directly to sediment or are mixed with soil before addition. A considerable sampling scheme may be necessary to overcome the spatial variability in concentrations in the sediment.

63. When applying the toxicological approach quantifying exposure may be done by taking three or four samples during one or two half-lives of the test substance in the relevant compartment(s), but at least 5 samples before 90% of the substance tested has disappeared, as long as such sampling is practical. This is done to estimate the temporal pattern of exposure, which may be employed to establish the exposure concentrations used for deriving a NOEC or an ECx. Subsequently monitoring is continued at a reduced frequency until concentrations fall below the level of biological concern, derived from lower-tier tests. Analysis of water is most commonly done; sediment analysis may also be important if the chemical has a high partition coefficient, degrades slowly in the sediment or is suspected to be toxic to benthic organisms. When applying the simulation type of approach degradation or dissipation or both of the test substance may also be followed during the test to establish the temporal pattern of exposure.

64. Depending on how much information is available before a microcosm or mesocosm study has begun, it may be necessary to obtain information about a) routes and rates of degradation b) chemical partitioning between water and sediment and between dissolved and particulate phases and c) uptake of test chemical by fish and other biota. Sampling should be done dependent on which aspects of fate are to be investigated. If the test substance bioaccumulates it may be necessary to take samples from macro-invertebrates, emerging insects or fish. It is recommended to attempt to reconstruct the mass balance of the

test chemical in order to account for all the test material introduced into the test system. This may be done with the aid of a radiotracer (Giddings, 1994), however this is not a requirement.

65. If possible, extraction of residues from water samples should be initiated immediately by the addition of a suitable solvent, when laboratory studies indicate rapid transformation of the test material in natural waters. Otherwise, samples should be refrigerated and extracted as soon as possible. Sediment samples should be frozen immediately. If there is a need to check stability of test substance during storage, water and sediment samples from untreated areas should be fortified with analytical standards, stored and analysed in the same manner as samples from treated areas. See also ISO guideline ISO/DIS 10381-6 (1993).

DATA AND REPORTING

Data handling and statistical analysis

66. Univariate analytical methods, such as ANOVA or regression analysis (or a combination of the two, e.g. William's test) are best suited to investigate parametric data on effects at the population level of one species or taxon. The power of these methods to detect differences from the control response, should be stated (Liber *et al*., 1992).

67. Multivariate analysis is appropriate for describing effects at the community level and can also be employed to indicate which taxa are particularly sensitive to treatment and would warrant specific univariate analysis. One method of multivariate analysis is based on the construction of Principal Response Curves (van den Brink & ter Braak, 1998, 1999) in which canonical coefficients for the identified taxa are plotted against time. This analysis takes into consideration the separate variances between replicates, between time-points and between treatments, thereby allowing clear representation of treatment effects in isolation. This pictorial evaluation of treatment effects can then be converted to an $NOEC_{community}$ with statistical significance using Monte Carlo permutation tests. Another method for evaluation of effects is the use of ecological models. Irrespective of the analytical tools applied for detecting the differences, the power of them may be considered. If the simulation type of approach was used than the effect or the no effect level should be related to loading.

68. Another method for evaluating microcosm or mesocosm studies is by calculating diversity or similarity indices. Analogous to the multivariate method described above also non-parametric multivariate methods may be used for the analysis of communities. Similarity indexes such as the Bray – Curtis similarity index (Bray & Curtis, 1957) may be coordinated for pictorial evaluation and statistically analysed using Monte Carlo permutation tests effectively separating time and treatment effects (Clarke, 1993; Clarke & Warwick 1994 & 2001). Whichever method, or combination of methods of analysis is to be applied, it is essential to build the analytical technique into the design of the study at the outset rather than to search for an appropriate method after the data have been generated.

69. It is recommended to analyse data from microcosm or mesocosm tests both by means of univariate and multivariate techniques to allow an evaluation at the population and community levels. It may then turn out that some NOEC's from individual species may be lower than the community NOEC. In that case, the ecological role and specific characteristics of that or those species and other related species should be considered.

70. It is recommended that the performer of a meso/microcosm study spend time to become familiar with the range of methods available for analysis before embarking on any one method. Sparks *et al*, 1999 provides a comprehensive overview of the range of ordination techniques, methods for the multivariate analysis of grouped data and methods that are available to examine relationships between groups of data. Further recommended literature includes: Chatfield and Collins, 1980; Digby and Kempton, 1987; Krzanowski, 1988 and Manly, 1986; Sokal and Rohlf, 1995; Sparks, 2000.

Appropriate levels of taxonomic resolution

71. Those taxa that are most sensitive should be identified to species (where practicable). Species-level identification is recommended for other taxa where practicable, since it may permit effective use of multivariate statistical approaches for more powerful analysis of community structure. The level of taxonomic analysis therefore depends on the objectives of the study. Univariate statistics may not be sufficiently powerful to detect differences among groups of rare organisms. In this case, data may be aggregated into larger taxonomic groups before analysis (Giddings *et al*. 2002).

72. Other groups of organisms that are identified as less sensitive in lower tiers (e.g. algae in tests with insecticides) may be monitored less intensively (e.g. at a lower level of taxonomic resolution such as family) or at a community level (e.g. chlorophyll *a* concentrations), although the ideal is species-level identification (There is always the chance of missed opportunities to identify new species entering the systems, if they are grouped, for example, as green algae. Furthermore, there is always the potential problem that within an "insensitive" group there may be a "sensitive" species.). Identification of organisms should be possible with available taxonomic keys and without breeding larval or nymphal forms through to older stages. Subsamples from abundant species (e.g. chironomids) may be analysed in detail to reduce counting labour and the proportion subsampled should be reported. It may become possible to analyse effects more rapidly in samples taken from microcosms by using flowcytometry for algae and image analysis for invertebrates. Currently such non-taxonomic methods are difficult to use because the significance of any measured change is unknown.

Structural versus functional endpoints

73. In general it can be said that conservation of function of an ecosystem is more robust than conservation of structure. However, when testing photosynthesis inhibiting herbicides for example functional endpoints (e.g. oxygen levels) may be more sensitive than structural ones (e.g. densities of algae and biomass of waterplants). In addition functional endpoints are integrative and their measurement can provide indications of the severity of impacts and consequences for the whole ecosystem. It is therefore important to evaluate structural effects in relation to the loss/maintenance of ecosystem function. It should also be noted that structural effects are important in themselves from the viewpoint of the maintenance of the biodiversity.

Reporting requirements

74. The final report should give a full and comprehensive description of the study, including its objectives, design and results. Along with description of the analytical and statistical techniques employed, the following data should also be reported, depending the study approach and on the objective of the study:

Information on test substance and relevant metabolites:

- identification, including chemical name and CAS number;

- batch or lot number;
- identification and levels of impurities;
- chemical stability under the conditions of the test;
- volatility;
- specific radioactivity and labeling positions (of appropriate);
- method for analysis of test substance and transformation products including limits of analytical detection/quantification;
- physico-chemical properties of the test substance, partition coefficients, rates of hydrolysis, photolysis, etc.;

Test systems:

- Description of test systems, location, history, dimensions, construction materials, general watershed characteristics, etc.;
- water levels and circulation;
- water quality: description of the chemical/physical parameters of the water used in the test system;
- colonization and introduction of biota;
- sediment characterization(with brief description of the water body where the sediment was extracted from);
- description of variation between replicates;

Experimental design and measured data:

- treatment regime: dosing regime, duration, frequency, loading rates, preparation of application solutions, application of test substance, etc.;
- sampling and analysis, residue monitoring results, analytical method;
- meteorological records;
- Physico-chemical water measurements (temperature, oxygen saturation, pH, etc);
- Sampling methods and taxonomic identification methods used;
- phytoplankton: chlorophyll-a; total cell density; abundance of individual dominant taxa; taxa (preferably species) richness, biomass;
- periphyton: chlorophyll-a; total cell density; density of dominant species; species richness, biomass;

- zooplankton: total density per unit volume; total density of dominant orders (Cladocera, Rotifera and Copepoda); species abundance; taxa richness, biomass;

- macrophytes: biomass, species composition and % surface covering of individual plants;

- emergent insects: total number emerging per unit time; abundance of individual dominant taxa; taxa richness; biomass; density; life stages;

- benthic macroinvertebrates: total density per unit area; species richness, abundance of individual dominant species; life stages;

- fish: total biomass at test termination; individual fish weights and lengths for adults or marked juveniles; condition index; general behaviour; gross pathology; fecundity, if necessary;

- possible effects of functional parameters (e.g. production, growth, activity or predation);

- possible data regarding adaptation and development of tolerance in organisms, species and/or communities;

Data evaluation:

- Endpoints of the study;

- Toxicity estimates (e.g. NOEC, ECx values), together with description of statistical methods used and discussion of statistical power;

- Treatment effects with univariate techniques;

- Treatment effect with multivariate techniques;

- Treatment effects with Similarity and Diversity Indices;

- Graphic presentation of results;

- Ecological significance of observed effects, with scientific rationale;

- Recovery of populations (observed or inferred), with a discussion of relevance to natural recovery processes; and

- The statistically derived NOEC or ECx. If another NOEC or ECx is regarded as the ecologically relevant result, then a scientific justification should be provided.

LITERATURE

ASTM 2004: Standard guide E1841-04 for conducting renewal phytotoxicity tests with freshwater emergent macrophytes. American Society for Testing and Materials, West Conshoken, PA. USA.

BEYERS, R.J. & ODUM, H.T. (1993) Ecological Microcosms. Springer-Verlag, New York, NY.

BRAY, J.R., CURTIS, J.T. (1957): "An ordination of the upland forest communities of Southern Wisconsin." Ecol. Monogr. 46: 327-354.

BROCK, T.C.M. and BUDDE, J.B. (1994). On the choice of Structural Parameters and Endpoints to Indicate Responses of Freshwater Ecosystems to Pesticide Stress. In:Freshwater Field Tests for Hazard Assessment of Chemicals. Eds. Hill, IR; Heimbach,.F.; Leeuwangh, P., Matthiessen, P.

BROCK, T.C.M., LAHR, J. & VAN DEN BRINK, P.J. (2000a) Ecological risks of pesticides in freshwater ecosystems. Part 1: Herbicides. Alterra-rapport 088, Wageningen, The Netherlands (can be obtained via www.pesticides.alterra.nl)

BROCK, T.C.M., VAN WIJNGAARDEN, R.P.A. & VAN GEEST, G.J. (2000b) Ecological risks of pesticides in freshwater ecosystems. Part 2: Insecticides. Alterra-rapport 089, Wageningen, The Netherlands (can be obtained via www.pesticides.alterra.nl)

BURNS, L.A., CLINE, D.M., and LASSITER, R.R. (1982). Exposure Analysis Modeling System (EXAMS): User Manual and System Documentation. EPA/600/3-82/0231 U.S. Environmental Protection Agency, Environmental Research Laboratory, Athens, GA.

CAIRNS, J., Jr., ED (1985) Multispecies Toxicity Testing. Pergamon Press, Elmsford, NY.

CAIRNS, J., Jr, ED (1986) Community Toxicity Testing. ASTM, Special Technical Pub 920, Philadelphia, PA.

CAIRNS, J., Jr & CHERRY, D.S. (1993) Freshwater Multi-species Test Systems. In Handbook of Ecotoxicology, Volume One. P. Calow (ed.) pp. 101-116, Blackwell Scientific Pub, London, UK.

CARSEL, R.F., SMITH, C.N., MULKEY, L-.A., DEAN, J.D. and JOWISE, P. (1984). User's manual for the Pesticide Root Zone Model (PRZM). Release 1. U.S. Environmental Protection Agency, EPA-600/384-109. U.S. Government Printing Office, Washington, DC.

CAMPBELL, P.J.; ARNOLD D.J.S; BROCK, T.C.M.; GRANDY, N.C.; HEGER, W.; HEIMBACH, F.; MAUND, S.J.; STRELOKE, M. Eds., (1999) Guidance Document on Higher-tier Aquatic Risk Assessment for Pexticides (HARAP). A SETAC-Europe Publication.

CHAPMAN P.F. & MAUND SJ. 1996. Considerations for the experimental design of aquatic mesocosm and microcosm studies. In: Techniques in Aquatic Toxicology. GK Ostrander (ed.) pp 657-673, CRC Lewis Publishers, Florida).

CHATFIELD, C, COLLINS AJ. 1980. Interoduction to Multivariate Analysis. Chapman & Hall, London, UK.

CLARK, J.R. & CRIPE, G.R. (1993) Marine and Estuarine Multi-species Test Systems. In Handbook of Ecotoxicology, Volume One. P. Calow (ed.) pp. 227-247, Blackwell Scientific Pub, London, UK.

CLARKE, K.R. (1993): "Non-parametric multivariate analysis of changes in community structure", Aust. J. Ecol. 18: 117-143.

CLARKE, K.R., WARWICK, R.M. (1994 & 2001): "Change in Marine Communities: An Approach to Statistical Analysis and Interpretation." 1st edition: Plymouth Marine Laboratory, Plymouth, UK, 144pp. 2nd edition: PRIMER-E, Plymouth, UK, 172pp

COMMITTEE TO REVIEW METHODS FOR ECOTOXICOLOGY. (1981) Testing for the Effects of Chemicals on Ecosystems. National Academy of Science Press, Washington, DC

COOPER, S.D. & BARMUTA, L.A. (1993) Field Experiments in Biomonitoring. In Freshwater Biomonitoring and Benthic Macroinvertebrates. D.M. Rosenberg, V.H. Resh (eds.) pp. 399-441, Chapman & Hall, New York, NY.

CROSSLAND, N.O. and LA POINT, T.W. (1992). The design of mesocosm experiments. Environmental Toxicology and Chemistry 11, 1-4.

DAY, K.E. (1989). Acute, chronic, and sublethal effects of synthetic pyrethroids on freshwater zooplankton. Environmental Toxicology and Chemistry 8:411-416.

DENOYELLES, F.,. W.D. KETTLE, C.H. FROMM, M.F. MOFFET, S.L. DEWY. (1989). Use of experimental ponds to assess the effects of a pesticide-on the aquatic environment. In J. R. Voshell, Jr. (ed.) Using Mesocosms to Assess the Aquatic Ecological Risk of Pesticides: Theory and Practice. Misc. Publications of the Entomological Society of America 75: 4-156.

DIGBY PGN, Kempton RA, 1987. Multivariate Analysis of Ecotoxicological Communities. Chapman and Hall, London, UK.

ECETOC. (1997) The Value of Aquatic Model Ecosystem Studies in Ecotoxicology. Technical Report No. 73. European Centre for Ecotoxicology and Toxiclogy of Chemicals, Brussels, Belgium.

ECOFRAM Aquatic report. In prep.

ECOFRAM Terrestrial report In prep.

EWOFFT (EUROPEAN WORKSHOP ON FRESHWATER FIELD TESTS). (1992). Summary and Recommendations of a meeting held at Potsdam, Germany; 25-26 June, 1992.

FAIRCHILD, J.F., LA POINT, T.W., ZAJICEK, J.L., NELSON, M.K, DWYER, F-J., LOVELY, P.A. (1992). Population-, community-, and ecosystem-level responses of aquatic mesocosms to pulsed doses of a pyrethroid insecticide. Environmental Toxicology and Chemistry 11: 115-129.

FOCUS (Forum for the Co-ordination of Pesticide Fate Models and their Use)., 1997. Soil persistence models and EU registration. Final report.

FOCUS 1997. Surface Water Models and EU registration of Plant Protection Products. Final report.

GIDDINGS, J.M. (1993). Sampling and analysis of the test chemical in microcosms and mesocosms. In: Hill, I.R., Heimbach, F., Leeuwangh, P., Matthiessen, P. (Eds), Freshwater Field Tests for Hazard Assessment of Chemicals. CRC Press, pp 151-158.

GIDDINGS, J.M., BROCK, T.C.M., HEGER, W., HEIMBACH, F., MAUND, S.J., NORMAN, S.M., RATTE, H.T., SCHÄFERS, C & STRELOKE, M. (2002). Community-Level Aquatic System Studies-Interpretation Criteria. Proceedings from the CLASSIC Workshop held at the Fraunhofer Institute, 30 May-2June, Smallenberg, Germany. SETAC PRESS, 44 pp

GRANEY R.L., GIESY, J.P, Jr., and DiTORO, D. (1989). Mesocosm experimental design strategies: advantages and disadvantages in ecological risk assessment. pp. 74-88, in VOSHELL, J.R., (1989). Using Mesocosms to Assess the Aquatic Ecological Risk of Pesticides: Theory and Practice. Miscellaneous Publication of the Entomological Society of America Nbr. 75. Entomological Society of America, Lanham, Maryland.

GRANEY, R.L., GIESY, J.P. & CLARK, J.R. (1995) Field Studies. In Fundamentals of Aquatic Toxicology, 2nd Ed. G.A. Rand (ed.) pp. 257-305, Taylor & Francis, Washington, DC.

GRANEY R.L., KENNEDY, J.H., and RODGERS, J.H., JR. (1993). Aquatic Mesocosm Studies in Ecological Risk Assessment. CRC Press, Boca Raton, Florida. 723 p.

HILL, I.R., Heimbach. F., Leeuwangh, P., Matthiessen, P. Eds. Freshwater Field Tests for Hazard Assessment of Chemicals. CRC Press, Inc. 1994.

HOWICK, G.L., GIDDINGS, J.M., deNOYELLES, F. FERRINGTON, L.C., JR., KETTLE, W.D., BAKER, D. (1992). Rapid establishment of test conditions and trophic-level interactions in 0.04-hectare earthen pond mesocosms. Environmental Toxicology and Chemistry 11: 107-114.

ISO/DIS 10381-6 (1993). Soil Quality – Sampling – Part 6: Guidance on the collection, handling and storage of soil for the assessment of aerobic microbial processes in the laboratory.

KERSTING K. (1994). Functional endpoints in field testing. In:Freshwater Field Tests for Hazard Assessment of Chemicals. Eds. Hill, IR; Heimbach,.F.; Leeuwangh, P., Matthiessen, P.

KERSTING, K & VAN WIJNGAARDEN, R.P.A. (1999). Effects of a pulsed treatment with the herbicide Afalon (active ingredient linuron) on macrophyte-dominated mesocosms I. Responses of ecosystem metabolism. Environ. Toxicol. Chem. 18: 2859-2865

KRZANOWSKI WJ. 1988. Principles of multivariate Analysis: A Users Perspective. Clarendon, Oxford, UK.

LEONARD, R. A., KNISEL, W.G., AND. D.A. STILL. (1987). GLEAMS: Groundwater Loading Effects of Agricultural Management Systems. Transactions of the American Society of Agricultural Engineers 30: 1403-1418.

LIBER, K., KAUSHIK N.K., SOLOMON, K.R., CAREY, J.H. (1992). Experimental designs for aquatic mesocosm studies: a comparison of the "ANOVA" and "Regression" design for assessing the impact of tetrachlorophenol on zooplankton populations in limnocorrals. Environmental Toxicology and Chemistry 11: 61-77.

LOZANO, S.J., O'HALLORAN, S.L., SARGENT, K.W., BRAZNER, J.C. (1992). Effects of esfenvalerate on aquatic organisms in littoral enclosures. Environmental Toxicology and chemistry 11: 35-47.

PARRISH, P.R.; DICKSON, K.L.; HAMELINK, J.L.; KIMERLE, R.A.; MACEK, K.J.; MAYER, F.L., Jr.; MOUNT, D.I. (1988). Aquatic toxicology: ten years in review and a look at the future. In W.J. Adams, G.A. Chapman, and W.G. Landis (Eds.) Aquatic Toxicology and Hazard Assessment: 10th Volume, STP971, American society for Testing and Materials, Philadelphia, Pennsylvania,.pp 7-25.

MANLY BFJ. 1986. Multivariate Statistical Methods – A Primer. Chapman & Hall. London, UK.

MATTHIESSEN, P. (1993). Guidelines for pesticide hazard assessment with freshwater field tests at the international level. In Heimbach, F. (ed.). "European Workshop on Freshwater Field Tests (EWOFFT). Lewis Publishers, Chelsea, Michigan.

OECD. (2006). Current Approaches in the Statistical Analysis of Ecotoxicity Data: A Guidance to Application, OECD Series on Testing and Assessment, No XX *(tbc)*.

SETAC. (1995). Procedures for assessing the environmental fate and ecotoxicity of pesticides. Ed. Mark R. Lynch. SETAC Europe, Brussels, Belgium, ISBN 90-5607-002-9, March 1995, 54pp

SETAC-EUROPE. (1991). Guidance Document on Testing Procedures for Pesticides in Freshwater Static Mesocosms. Report from a meeting held at Monks Wood Experimental Station, Huntingdon, UK; 3-4 July, 1991. 45 p.

SETAC/RESOLVE. (1991). Workshop on Aquatic Microcosms for Ecological Assessment of Pesticides. Report from a meeting held at Wintergreen Virginia, USA. October 1l, 1991.

SOKAL RR, ROHLF FJ, 1995. Biometry. WH Freeman, New York, NY, USA.

SPARKS, TH, SCOTT, WA, CLARKE, RT. 1999. Traditional multivariate techniques: Potential for use in Ecotoxicology. Env. Tox. & Chem. 18, No.2 pp 128-137.

SPARKS, TH (ed.). 2000. Statistics in ecotoxicology. John Wiley & Sons, LTD, Chichester, UK

STEPHEN, C.E, RODGERS, J.W. (1985). Advantages of using regression analysis to calculate results of'chronic toxicity tests, pp. 328-338. In R.C. Bahner & D.J. Hansen (Eds.) Aquatic Toxicology and Hazard Assessment: Eighth Symposium, ASTM STP 891. American Society for Testing and Materials, Philadelphia, Pennsylvania.

SUGIURA, K. (1992). A multispecies laboratory microcosm for screening ecotoxicological impacts on chemicals. Env. Tox, and Chem. 11: 1217-1226.

TOUART, L.W. (1988). Aquatic mesocosm tests to support pesticide registrations. Technical Guidance Document, Hazard Evaluation Division. U.S. Environmental Protection Agency, Washington DC. EPA/540/09-88-035.

VAN DEN BRINK, P.J., & TER BRAAK, C.F.J., 1998. Multivariate analyis of stress in experimental ecosystems by Principle Response Curves and similarity analysis. Aquatic Ecology 32: 163-178.

VAN DEN BRINK, P.J., & TER BRAAK, C.F.J., 1999. Principle Response Curves: analysis of time-dependent multivariate responses of a biological community under stress. Env. Tox. and Chem. 18: 138-148.

VAN WIJNGAARDEN, R.P.A.; VAN DEN BRINK, P.J.; CRUM, S.J.H.; OUDE VOSHAAR, J.H.; BROCK, T.C.M.; LEEUWANGH, P. (1996). Effects of the Insecticide Dursban® 4^E (Active Ingredient Chlorpyrifos) in outdoor experimental ditches: I. Comparison of Short-term Toxicity between the Laboratory and the Field. Env. Tox. and Chem., Vol. 15, No. 7, 00 1133-1142.

VOSHELL, J.R., (1989). Using Mesocosms to Assess the Aquatic Ecological Risk of Pesticides: Theory and Practice. Miscellaneous Publication of the Entomological Society of America. Nbr. 75. Entomological Society of America, Lanham, Maryland. 88 p.

WETZEL & LIKENS (1991). Limnological Analyses. New York: Springer-Verlag, 392 pp.

WILLIAM'S D.A., (1971) A test for differences between treatment means when several dose levels are compared with a zero dose control. (*Biometrics,* 27, 103-117).

WORLD WILDLIFE FUND/RESOLVE. (1992). Improving Aquatic Risk Assessment under FIFRA. Report of the Aquatic Effects Dialogue Group. RESOLVE, Washington, DC. 98 p.

Table 1: Description and Comparisons of Typical Experimental Systems: Simulated Freshwater Lentic Field Tests (Outdoor Microcosms and Mesocosms)

Characteristic/Parameter (Dependent on Study Objectives)	Outdoor Lentic Microcosm	Outdoor Lentic Mesocosm	Larger (Whole System) Lentic Mesocosm
Size (Volume)	10^{-3} to 10 m^3	1 to 10^4 m^3	10^3 to 10^8 m^3
Time (Temporal Scale)	Hours to several weeks or months	Days to many months	Weeks to several years
Container Examples	Glass, plastic, stainless steel, epoxy-resin, soil-lined vats, tubs, tanks, pools, concrete ponds	Small pond, enclosed portion of a larger pond or small lake (e.g., Lund rubber tubes, bags, cylinders, or liners), limnocorrals	Large earthen pond, small lake, larger enclosures
Similarity to Natural Ecosystems (Low to High)	Low to moderate	Moderate to high	High
Organisms Included	Primary producers (algae, periphyton), and invertebrate herbivores and consumers, usually no fish. May include macrophytes.	All types, may include macrophytes and fish.	All types, including macrophytes and fish.
Parameters Measured (At population and community levels)	Weather conditions, and various water quality parameters (e.g., pH, alkalinity, hardness, dissolved oxygen, temperature). Mortality, growth, reproduction, diversity, similarity, succession, abundance (numbers, biomass), and composition of organisms and populations. Primary production (photosynthesis and respiration), chemical fate (esp. uptake), and nutrient cycling. Recovery.	Same, with more focus on community level parameters. Recovery.	Same, and longer term community parameters such as succession of species, and multi-seasonal changes in populations, community grazing, competitive relationships, recovery.
Number of Replicates	Three or more	Several (Two or more)	One or two, but it may not be possible to establish true replicates due to size and complexity of system.
Number of Test Substance Concentrations	Five or more	Several (Three or more)	One or two
Water	Non-polluted, uncontaminated source.	Same	Whatever is present in the system is used

	Well-water, aged tap water, or water from a natural source.		and should be uncontaminated and unpolluted.
Sediments	Natural, uncontaminated, unpolluted source.	Same	Whatever is present in the system is used and should be uncontaminated and unpolluted.
Other Advantages (also see above)	Can study the effects and fate of chemical substances in a "semi-natural" field setting. Cost to perform test would be the least. Sampling somewhat easier to perform. Time to complete test could be short.	Same except system resembles more closely a natural outdoor lentic ecosytem. Cost to perform test would be moderate. Time to complete test could be short to moderate.	Same except this system resembles most closely a natural outdoor lentic ecosystem. Easier to establish a diverse and representative community. Less affected by short-term environmental variations. Less divergence from natural conditions. May be a good predictor of effects of test substances on natural systems.
Other Limitations (also see above)	More affected by short-term environmental variations; less affected by long-term environmental variations. System resembles less a natural outdoor, lentic ecosystem. Container edge effects may be present, and mixing lessened. Easier to diverge from natural conditions. More difficult to establish a diverse, and representative community. May be high variability between replicates.	Moderately affected by both short-term and long-term environmental variations. True replication difficult. Edge effects and divergence from natural conditions is less, but still may be present. There may be high variability between replicates.	More affected by long-term environmental variations. True replication not possible. Sampling is more difficult and complicated. No container edge effects. Time to complete test could be moderate to lengthy. High costs. Inability to apply all but minimal control on this system.
Some Examples	Woltering (In Cairns, 1985); Heimbach et al., Howick et al., Johnson et al. (In Graney et al., 1994); Draxl et al., Thielcke and Ratte (In Hill et al., 1994).	Hook et al. (In Cairns, 1986); Christman et al., Hill et al., Liber et al., Thompson et al. (In Graney et al., 1994).	Hook et al. (In Cairns, 1986), Brezonik et al., Schindler et al. (In Cairns and Cherry, 1993); Ferrington et al. (In Graney et al., 1994), Stephenson (In Hill et al., 1994).

Sources: Cairns, (1985, 1986), Cairns and Cherry (1993), Campbell et al. (1999), Cooper and Barmuta (1993), ECETOC (1997), Giddings et al. (2002), Graney et al. (1994), Hill et al. (1994), SETAC-EUROPE (1991),

Table 2: An example of some typical parameters measured in Microcosm or mesocosm studies and their frequency of measurement [1]

	Parameter	Suggested frequency
Water quality	Water level, pH, temp. DO, turbidity, conductivity, alkalinity, hardness, suspended solids, nutrients (dissolved C)	At least every two week intervals.
	Pesticides, heavy metals	At initiation of the test
Sediment quality	Pesticides, metals, particle size, ion exchange capacity, organic content, pH	At initiation of the test
Phytoplankton	Chlorophyll a /phaeophytin/dry wt., cell counts and additionally species-diversity for long-term tests. Cell density	At least every two weeks
Periphyton	Chlorophyl a + phaeophytin + dry wt; Cell counts	At least twice during the study
Macrophytes	Identify and estimate abundance visually (+ photographically) Biomass by fresh weight	At infrequent intervals during the peak growth period and at finish
Macro-invertebrates	Benthos Adult insects Artificial substrates + emergent insects Optional grabbing Identify to lowest practical taxon	Every two weeks; For adult insects; weekly during peak emergence time and at time of test substance application – during other periods less frequent sampling may be sufficient.
Zooplankton	Identify species if possible Density and biomass Record life stages	Weekly
Fish	Length/weight	At beginning and end
	Dead fish weighed and measured Gross pathology Sex/fecundity if relevant	At end
Residues (test substance concentration)	Test material + degradates	Frequency dependent on compound. More frequently in the beginning than at the end of the experiment.
Meteorology	Air temp, solar radiation, precipitation, windspeed	At appropriate intervals, on site if possible.

1. From "Guidelines for Pesticide Hazard Assessment with Freshwater Field tests at the International level", Matthiesen, P. In: Hill et al, 1994.