SCIENTIFIC OPINION

Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters[1]

EFSA Panel on Plant Protection Products and their Residues (PPR)[2,3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

EFSA's Panel on Plant Protection Products and their Residues (PPR) was tasked to revise the Guidance Document (GD) on Aquatic Ecotoxicology under Council Directive 91/414/EEC (SANCO/3268/2001 rev.4 (final), 17 October 2002). This Guidance of the PPR Panel is the first of three requested deliverables within this mandate. It has its focus on tiered acute and chronic effect assessment schemes with detailed guidance on tier 1 and higher tier effect assessments for aquatic organisms in edge-of-field surface waters and on proposals regarding how to link effects to exposure estimates. The exposure assessment methodology was not reviewed and it is assumed that the current FOCUS surface water exposure assessment methodology will continue to be used for exposure assessment at EU level. The current GD is intended to be used for authorisation of active substances at EU level as well as for plant protection products at Member State level. The effect assessment schemes in this GD allow for the derivation of regulatory acceptable concentrations (RACs) on the basis of two options: (1) the ecological threshold option (ETO), accepting negligible population effects only, and (2) the ecological recovery option (ERO), accepting some population-level effects if ecological recovery takes place within an acceptable time period. In the tiered effect assessment schemes, in principle, all tiers (1, 2 and 3) are able to address the ETO, while the model ecosystem approach (tier 3), under certain conditions, is able to also address the ERO. The GD provides the scientific background for the risk assessment to aquatic organisms in edge-of-field surface waters and is structured to give detailed guidance on all assessment steps. An executive summary joining all parts of the guidance and decision schemes in a concise way is provided and is intended to help applicants and regulatory authorities in day-to-day use.

KEY WORDS

pesticides, formulations, metabolites, ecotoxicology, aquatic organisms, specific protection goals, regulatory acceptable concentrations

---

[1] On request from EFSA, Question No EFSA-Q-2009-00001, adopted on 20 June 2013.

[2] Panel members: Alf Aagaard, Theo Brock, Ettore Capri, Sabine Duquesne, Metka Filipic, Antonio F. Hernandez-Jerez, Karen I. Hirsch-Ernst, Susanne Hougaard Bennekou, Michael Klein, Thomas Kuhl, Ryszard Laskowski, Matthias Liess, Alberto Mantovani, Colin Ockleford, Bernadette Ossendorp, Daniel Pickford, Robert Smith, Paulo Sousa, Ingvar Sundh, Aaldrik Tiktak, Ton Van Der Linden. Correspondence: pesticides.ppr@efsa.europa.eu.

[3] Acknowledgement: EFSA wishes to thank the members of the Working Group Aquatic Ecotoxicology: Alf Aagaard, Paulien Adriaanse, Jos Boesten, Theo Brock, Sabine Duquesne, Michael Klein, Matthias Liess, Robert Luttik, Paul Miller, Daniel Pickford, Aaldrik Tiktak, Jan Vanderborght, Lina Wendt-Rasch and EFSA staff: Stephanie Bopp, Maria Arena and Alessandra Caffi for the support provided to this scientific opinion.

Suggested citation: EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2013. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Journal 2013;11(7):3290, 186 pp. doi:10.2903/j.efsa.2013.3290.

Available online: www.efsa.europa.eu/efsajournal

SUMMARY

The European Food Safety Authority (EFSA) asked the Panel on Plant Protection Products and their Residues (PPR) to prepare a revision of the Guidance Document on Aquatic Ecotoxicology under Council Directive 91/414/EEC (SANCO/3268/2001 rev.4 (final), 17 October 2002; EC, 2002a). The PPR Panel was therefore tasked to prepare a revised Guidance Document and two Scientific Opinions. This Guidance of the PPR Panel is the first of these three requested deliverables as outlined in the Terms of Reference below. The revision of the former Guidance Document on Aquatic Ecotoxicology became necessary mainly due to (1) the entry in to force of the new Regulation (EC) No 1107/2009[4] on authorisation of plant protection products, (2) the revision of the related data requirements and (3) scientific developments. Stakeholders were consulted before the start of the revision process in a public consultation, as well as risk managers in a specific consultation, in October to December 2008. The revision of the Guidance Document on Aquatic Ecotoxicology was started in parallel to the revision of the Guidance Document on Terrestrial Ecotoxicology (SANCO/10329/2002, rev.2 final, 17.10.2002; EC, 2002b) to allow a harmonisation process. As a first step, the PPR Panel developed a framework for deriving specific protection goals (EFSA PPR Panel, 2010a). The approach outlined in that opinion was the starting point for the development of this updated Guidance Document on aquatic risk assessment (RA).

The aquatic RA is the combination of the exposure and the effect assessments and there is considerable interaction between these assessments. The focus of this Guidance Document (GD) is on a tiered effect assessment scheme with detailed guidance on tier 1 and higher tier effect assessments that are mainly based on experimental approaches (chapters 7–10). A scientific opinion on the state of the art of mechanistic effect modelling in the aquatic environment (e.g. toxicokinetic/toxicodynamic and population models) will be delivered later under this mandate. The effect assessment guidance is intended to be used for authorisation of active substances (a.s.) at EU level as well as for plant protection products at Member State level. Furthermore, the appropriate linking between exposure and effect assessment is described. The exposure assessment methodology was not reviewed and it is assumed that the current FOCUS surface water exposure assessment methodology will continue to be used for exposure assessment at EU level. Only a brief overview of the exposure assessment is included in this GD in chapter 6, for details reference is made to the related FOCUS surface water guidance (FOCUS, 2001).

The GD first describes the specific protection goals for aquatic organisms that need to be defined in order to develop an appropriate RA scheme. Proposed specific protection goals (SPGs) were discussed with risk managers in September to November 2012 and are described in chapter 5. The SPGs overall aim is to protect aquatic plants and animals at the population level in surface water. However, the SPG selected for aquatic vertebrates aims at protection at the individual level, so that mortality and suffering due to acute toxicity is avoided. As outlined in the PPR Panel opinion on SPGs (EFSA PPR Panel, 2010a), the exposure assessment goals also have to be defined in parallel to set the overall level of protection. Since the exposure assessment methodology was not revised in parallel to the effect assessment scheme, definitions for exposure assessment goals are not clear.

In this GD, the tiered effect assessment procedure and proposals on how to link effects to exposure estimates are focused on aquatic organisms living in the water column of edge-of-field surface waters. For these organisms, the concentration of the freely dissolved chemical is chosen as the ecotoxicologically relevant concentration (ERC). This GD also presents the tier 1 effect assessment procedure for sediment-dwelling organisms when based on water-sediment toxicity tests. More information for sediment-dwelling organisms will be provided in an opinion on the effect assessment for plant protection products on sediment organisms in edge-of-field surface water to be delivered next under this mandate.

[4] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L 309/1, 24.11.2009, pp. 1–50.

To protect populations of aquatic organisms, effect assessment schemes are developed that allow for the derivation of regulatory acceptable concentrations (RACs) on the basis of two options: (1) the ecological threshold option (ETO), accepting negligible population effects only, and (2) the ecological recovery option (ERO), accepting some population-level effects if ecological recovery takes place within an acceptable time period. In the tiered acute and chronic effect assessment schemes, in principle, all tiers (1, 2 and 3) are able to address the ETO, while the model ecosystem approach (tier 3), under certain conditions, is able to also address the ERO. The ETO from tier 3 is particularly relevant as it is more likely to assure an adequate level of protection, not only for the application of a single plant protection product (PPP), but also in view of the application of (non-approved) tank-mixtures and serial PPP applications during the growing season. It thus may better address issues of the 'uniform principles' as laid down in Regulation (EC) No 546/2011[5] that requires that Member States base their authorisation decision on the 'proposed conditions for the use of the plant protection product' and furthermore the standard data requirements for PPP do request: 'any information on potentially unacceptable effects of the plant protection product on the environment, on plants and plant products shall be included as well as known and expected cumulative and synergistic effects'.

The GD is structured to give detailed guidance and provide relevant scientific background information for each tier in the respective Chapters, which all end with a section on how to derive RACs and how to perform the RA, including decision schemes. Chapter 7 describes the tier 1 effect assessment based on the revised data requirements. Chapter 8 addresses refinement options based on additional species tested, that is, the Geomean approach and the species sensitivity distribution approach. Chapter 9 addresses higher tier options based on refined exposure laboratory and model ecosystem approaches. This includes guidance on selecting the appropriate refined exposure profiles, on refined exposure laboratory tests, and on designing and evaluating model ecosystem (micro-/mesocosm) studies. Chapter 10 contains detailed guidance on the possible use of non-testing methods, effect assessment for metabolites, and assessment for formulations containing more than one a.s.. Chapter 11 addresses other relevant related issues. Chapter 12 provides guidance on addressing the uncertainties in the assessment.

Chapter 2 provides an executive summary that joins all guidance and decision schemes in a concise way without the detailed scientific background. This is intended to be helpful to applicants and regulatory authorities providing an overview for day-to-day use.

The guidance was developed based on experience with currently known or approved a.s. and plant protection products. When using this guidance, it should be always checked whether the proposed schemes are appropriate for a.s. with a new mode of action.

[5] Regulation (EC) No 546/2011 of 10 June 2011 implementing Regulation (EC) No 1107/2009 of the European Parliament and of the Council as regards uniform principles for evaluation and authorisation of plant protection products.

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

Member States' competent authorities were requested by the Director of Sciences of the European Food Safety Authority (EFSA) on 3 July 2006 *via* the Standing Committee on the Food Chain and Animal Health, to send EFSA a priority list of existing Guidance Documents to be revised and proposals for development of new ones. Answers were received from 15 Member States.

Regarding the revision of the Guidance Document on Aquatic Ecotoxicology (SANCO/3268/2001, rev. 4 final, 17 October 2002), five detailed requests were received (FI, DE, NL, DK, SE) highlighting the importance of liaising with the revision of Annex II and Annex III.

In 2006 and 2007, EFSA has issued six opinions on the Annexes II and III, two of which related to the ecotoxicological studies (EFSA, 2007a) and the fate and behaviour in the environment (EFSA, 2007b). The rapporteur (UK) has taken these opinions on board in the revision of the Annexes, which are currently with the Commission. It should be considered to generally revise the structure and content of the available Guidance Documents.

Member States highlighted the following issues as being particularly important:

- More clarity regarding the data requirements for substances expected to be endocrine disrupters is needed;
- More guidance should be provided regarding the use of $FOCUS_{SW}$ modelling, e.g. on input parameters or the use of Step 4;
- Need for revision in particular with regard to the protection level in adjacent small ditches and main watercourses (in line with the requirements of the Water Framework Directive);
- More integrated development of the assessment of exposure modelling and effects;
- Conceptual consistency between higher tier assessments in aquatic and terrestrial ecotoxicology needed;
- More guidance regarding the assessment of higher tier aquatic studies (assessment of addition of sediments, assessment of quality and quantity of mesocosm studies, assessment of ecotoxicological field studies, trigger levels for higher tier studies);
- Harmonised endpoints for authorisation of plant protection products needed;
- A clear and transparent relationship with the Water Framework Directive is wished for.

The EFSA *PRAPeR Unit* emphasised that the aquatic GD needs to be updated regarding the long-term RA to take account of the new exposure data that are the outcome of the FOCUS models. The interaction between exposure and effects needs some more guidance. Of course also possible new data requirements in the new regulation that will replace Council Directive 91/414/EEC need to be taken up in the existing GD.

Relevant topics and scientific principles of already existing scientific opinions elaborated by the PPR Panel will also be incorporated into the revised Guidance Document. Further, on-going work in other fora, pertinent to the GD will be closely monitored and taken into account where relevant.

The public was consulted on the existing GD in October – December 2008 and comments and ideas for the revision by stakeholders will be taken into account during the process. Also comments from a risk manager survey performed October – December 2008 are considered. Furthermore, the activity performed under EFSA-Q-2009–00861 to develop specific protection goals will be used as input to this updated mandate.

TERMS OF REFERENCE AS PROVIDED BY EFSA

EFSA tasks its Scientific Panel on Plant Protection Products and their Residues (PPR Panel) to prepare a revision of the Guidance Document on Aquatic Ecotoxicology under Council Directive 91/414/EEC (SANCO/3268/2001 rev.4 (final), 17 October 2002).

The PPR Panel is asked to develop a Guidance Document and two Scientific Opinions, as summarised below:

1. **Guidance Document on tiered risk assessment for aquatic organisms in edge-of-field surface waters** (by July 2013).

In particular, the following issues need to be addressed:

- Update the current guidance in view of the new Regulation (EC) No 1107/2009
- Update the current guidance in view of the revised data requirements to Regulation (EC) No 1107/2009
- Develop guidance on first tier aquatic effect assessment
- Develop guidance on higher tier aquatic effect Assessment (based on laboratory studies and model ecosystem studies, guidance on design and evaluation of higher tier studies)
- Guidance on appropriate linking of aquatic exposure and effect assessment

This PPR Panel Guidance should be subject to a Public Consultation.

2. **Scientific Opinion of the PPR Panel on the effect assessment for pesticides on sediment organisms in edge-of-field surface waters** (2 years after acceptance of the revised mandate, i.e. October 2014)

A scientific opinion will be provided that describes the state of the art of effect assessment for sediment organisms.

In particular the following issues will be addressed:

- Identification of standard test species
- Use of the geometric mean approach when toxicity data for a limited number of additional test species are available
- Use of Species Sensitivity Distribution approach for sediment organisms
- Use of the model ecosystem approach for sediment organisms
- Defining the ecotoxicologically relevant concentrations (ERCs) for acute and chronic risk assessment

3. **Scientific Opinion on the state of mechanistic effect modelling approaches for regulatory risk assessment of pesticides for aquatic organisms** (3.5 years after acceptance of the revised mandate, i.e. April 2016)

A scientific opinion will be provided that describes the state of the art of mechanistic effect modelling in the aquatic environment.

In particular the following state of the art of the following types of models will be addressed (for all aquatic water column and sediment dwelling organisms):

- Describe regulatory questions that can be addressed by effect modelling
- Describe model parameters that need to be included in relevant models and that need to be checked in evaluating the acceptability of effect models
- Describe available effect models for aquatic organisms, in particular
    - Toxicokinetic/toxicodynamic models
    - Mechanistic population models
    - Mechanistic food web models
        - Secondary poisoning
        - Ecosystem models representative for ditches, ponds and streams
- Selection of focal species
- Development of ecological scenarios that can be linked to the regulatory defined water bodies in the climatic zones of Europe

This Guidance Document addresses the first part of the Terms of Reference, the two scientific opinions outlined above will follow later.

ASSESSMENT

1. Reading guidance

The Guidance Document (GD) is structured to give detailed guidance and provide relevant scientific background information for each tier in the respective Chapters, which all end with a section on how to derive regulatory acceptable concentrations (RACs) and how to perform the risk assessment (RA), including decision schemes. Chapter 3 describes the focus and restrictions of the GD. Chapter 4 introduces the tiered approach, terminology and the linking of exposure to effects. Chapter 5 describes the specific protection goals. Chapter 6 addresses the exposure assessment according to the current FOCUS surface water methodology. Chapter 7 describes the tier 1 effect assessment based on the revised data requirements. Chapter 8 addresses refinement options based on additional species tested, that is, the Geomean approach and the species sensitivity distribution approach. Chapter 9 addresses higher tier options based on refined exposure laboratory and model ecosystem approaches. This includes guidance on selecting the appropriate refined exposure profiles, on refined exposure laboratory tests, and on designing and evaluating model ecosystem (micro-/mesocosm) studies. Chapter 10 contains detailed guidance on the possible use of non-testing methods, effect assessment for metabolites, and effect assessment for formulations containing more than one active substance (a.s.). Chapter 11 addresses other relevant related issues. Chapter 12 provides guidance on addressing the uncertainties in the assessment.

Chapter 2 provides an executive summary that joins all guidance and decision schemes in a concise way without the detailed scientific background. This is intended to be helpful to applicants and regulatory authorities in day-to-day use. Appendix H gives examples of where this new GD was applied to perform the RA for three imaginary a.s..

2. Executive summary

This executive summary joins all guidance and decision schemes of the later chapters in this document in a concise way without the detailed scientific background. This is intended to be helpful to applicants and regulatory authorities in day-to-day use. For the details, references to the relevant section(s) with more detailed background information and explanation are included.

2.1. Aquatic risks due to toxicity

2.1.1. Introduction

The tiered effect assessment procedure, and proposals for how to link effect to exposure estimates, presented in this revised aquatic guidance document (AGD), focus on aquatic organisms living in the water column of permanent edge-of-field surface waters. For these organisms, the concentration of the freely dissolved chemical (hence not including chemical sorbed, for example, on suspended matter or sediment) is chosen as the ecotoxicologically relevant concentration (ERC). The AGD also presents the tier 1 effect assessment procedure for sediment-dwelling organisms based on water-sediment toxicity tests.

The aquatic RA combines exposure and the effect assessments and there is considerable interaction between these assessments. This guidance document assumes that the current exposure assessment procedure (FOCUS surface water scenarios and models) continues to be used at the EU level for approval of a.s., and does not include further guidance for the exposure assessment. To date, the PPR Panel has not evaluated the current exposure assessment procedure. The overall level of protection of aquatic organisms is determined by the combination of the specific protection goals (SPGs) for the organisms and the exposure assessment goals. Since the exposure assessment methodology was not revised in parallel to the effect assessment scheme, the overall level of protection remains unclear. The PPR Panel advises to critically evaluate and improve the surface water exposure assessment in the future.

Two distinct effect assessment schemes are elaborated that respectively start with the tier 1 acute and the tier 1 chronic toxicity data set, respectively (see Figure 1 for a schematic overview; for a more detailed description see sections 4.3 and 4.4). The acute and chronic effect assessment schemes address the same SPGs. The overall aim of these SPGs is to protect aquatic plants and animals at the population level in surface water. However, the SPG selected for aquatic vertebrates aims at protection at the individual level, so that mortality and suffering due to acute toxicity is avoided (see section 5.5 for more details).

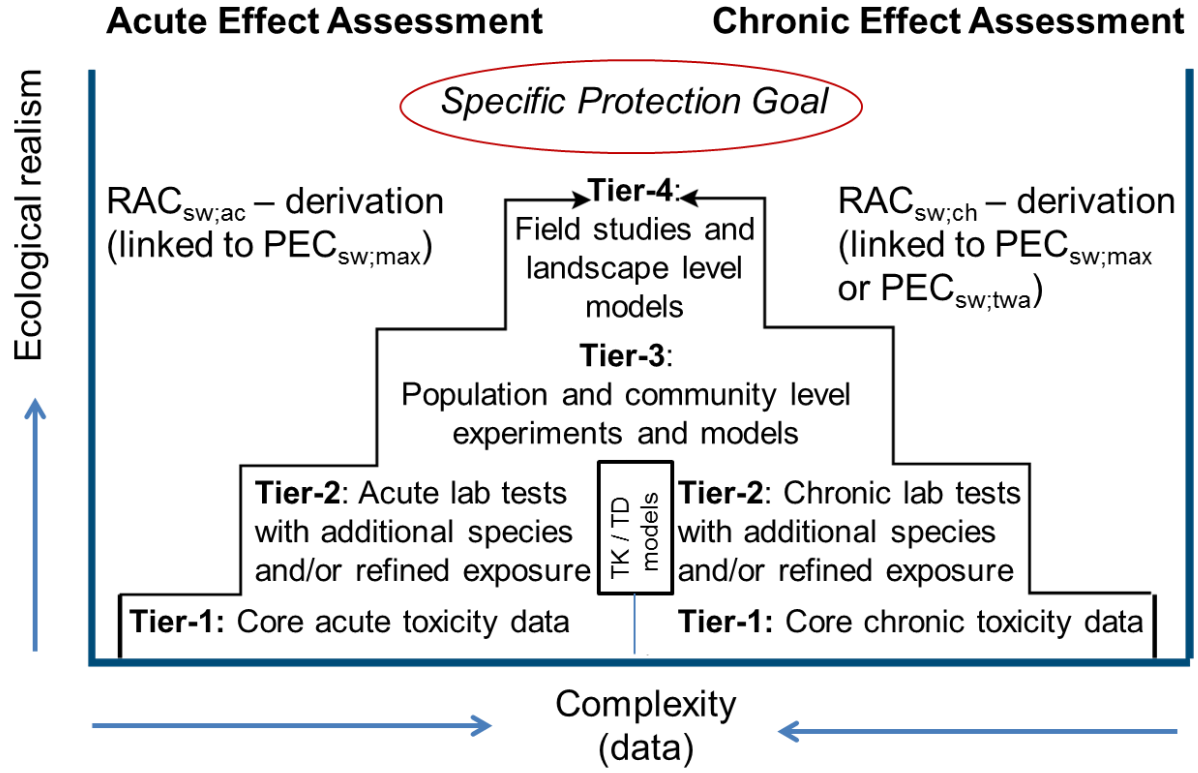


Figure 1: Schematic presentation of the tiered approach within the acute (left part) and chronic (right part) effect assessment for plant protection products (PPPs). For each PPP, both the acute and chronic effects/risks have to be assessed. The tier 1 and tier 2 effects assessments are based on single species laboratory toxicity tests, but to better address risks of time-variable exposures the tier 2 assessment may be complemented with toxicokinetic/toxicodynamic (TK/TD) models. Tier 3 (population- and community-level experiments and models) and tier 4 (field studies and landscape-level models) may concern a combination of experimental data and modelling to assess population- and/or community-level responses (e.g. recovery, indirect effects) at relevant spatio-temporal scales. All models included in such a tiered approach need to be properly tested and fulfil required quality criteria.

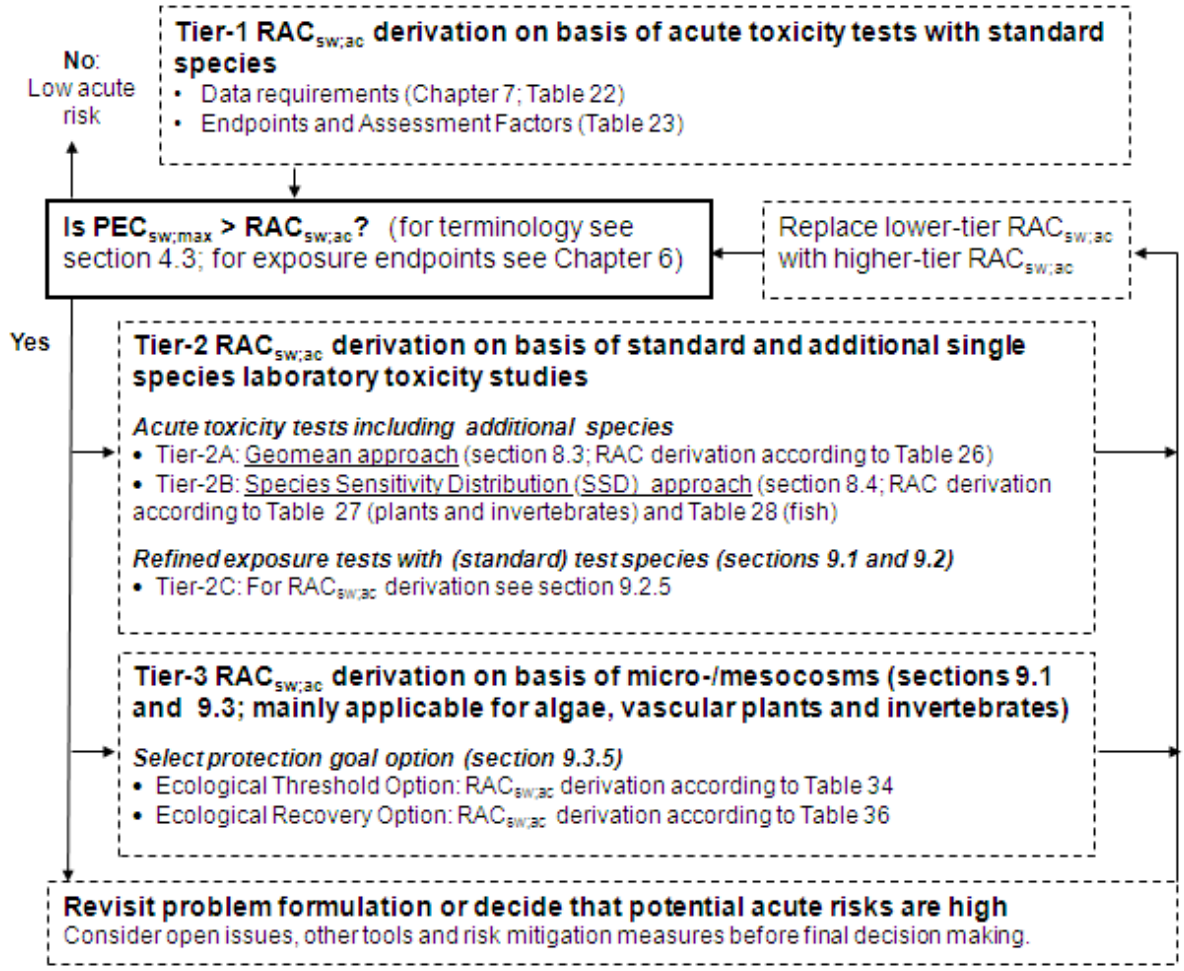
Although in the effect assessment schemes presented in Figure 1 also modelling procedures are mentioned, this GD mainly updates guidance on experimental approaches (tiers 1, 2 and 3) to derive Regulatory Acceptable Concentrations (RACs). Effect models in the aquatic RA of PPPs will be the subject of a future scientific opinion of the PPR Panel of EFSA.

To protect populations of aquatic organisms, effect assessment schemes are developed that allow the derivation of RACs on the basis of two options: (1) The ecological threshold option (ETO), accepting negligible population effects only, and (2) the ecological recovery option (ERO), accepting some population-level effects if ecological recovery takes place within an acceptable time period (see section 5.5). In the tiered acute and chronic effect assessment schemes presented in Figure 1, in principle all tiers are able to address the ETO, while the model ecosystem approach (tier 3), under

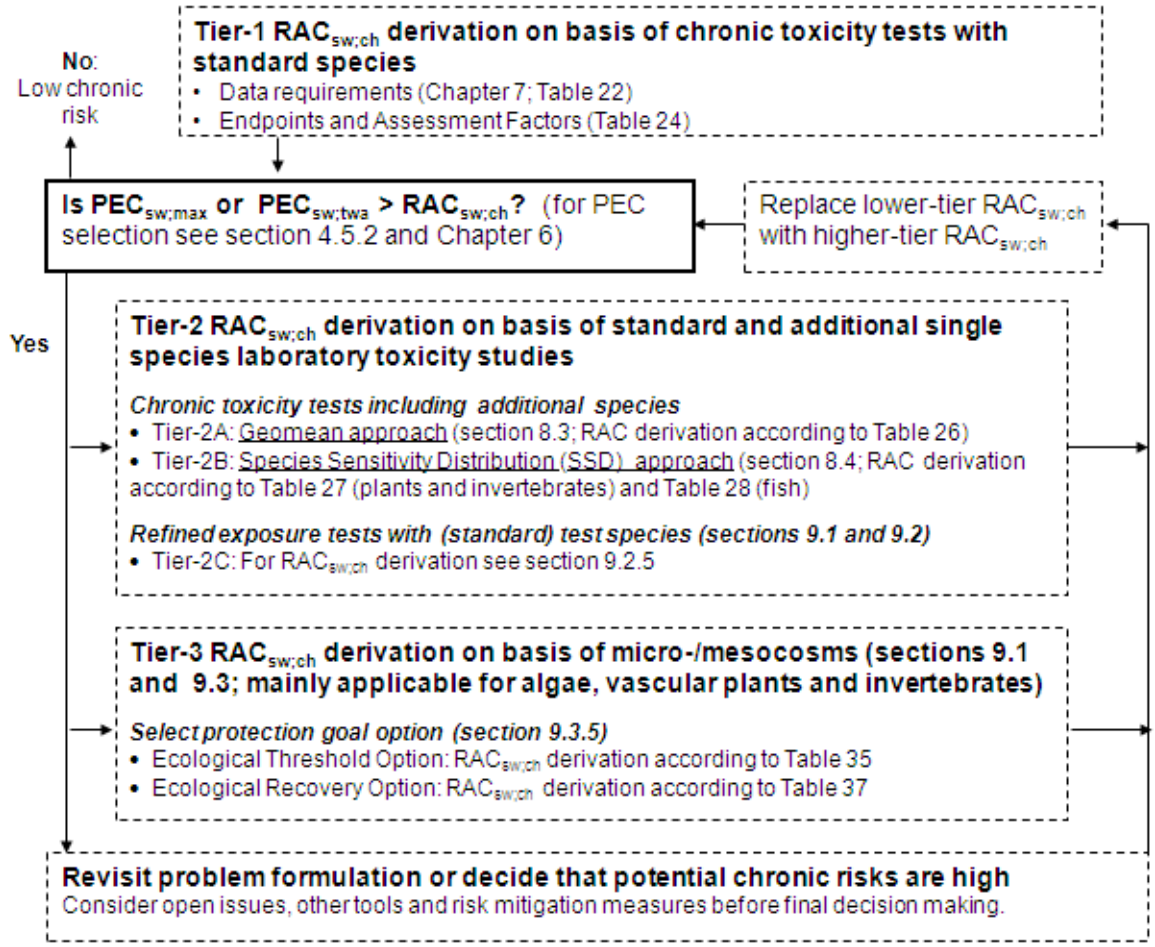
certain conditions (e.g. possibility to extrapolate observed responses to potential vulnerable species), is able to also address the ERO.

2.1.2. Summary flow charts for acute and chronic effect/risk assessment

The basic flow chart for the tiered procedure to assess acute effects/risks of predicted PPP exposure in edge-of-field surface water is presented in Decision scheme A. The procedure for chronic effect/risk assessment is presented in Decision scheme B. In these schemes, the specific chapters/sections of the GD are mentioned where detailed guidance for decision making, as well as information on the scientific rationale, can be found. Note that in the registration procedure for each a.s. both the acute and chronic effect/risk assessment procedures have to be performed and that a higher tier RAC may be valid for specific groups of organisms only (e.g. invertebrates and/or primary producers when derived from a mesocosm test). Consequently, it should always be checked whether the higher tier RAC is valid for non-tested groups of organisms to be protected (in an iterative process, by comparing results of RACs generated by higher tier RACs assessments with those generated in lower tier RACs).



Decision scheme A: Summary flow chart for acute effect/risk assessment of toxicity. Chapters/sections of this aquatic GD are mentioned where detailed guidance for decision making, as well as its scientific underpinning, can be found.



Decision scheme B: Summary flow chart for chronic effect/risk assessment of toxicity. Chapters/sections of the aquatic GD are mentioned where detailed guidance for decision making, as well as its scientific underpinning, can be found.

In the acute (ac) RA, the $RAC_{sw;ac}$ (RAC in surface water (SW) for adverse effects of pesticide exposure occurring within a relatively short period after exposure) is always compared with the $PEC_{sw;max}$ (the maximum predicted environmental concentration (PEC) in surface water) derived from the predicted exposure profile. In the chronic (ch) RA, the $RAC_{sw;ch}$ (RAC in surface water for adverse effects of pesticide exposure that develop slowly and/or have a long-lasting course and that are caused by short- or long-term exposure) is in the first instance compared with the $PEC_{sw;max}$, and under certain conditions with a $PEC_{sw;twa}$ (the predicted time-weighted average (TWA) concentration in surface water). A decision scheme on when to use the $PEC_{sw;max}$ or the $PEC_{sw;twa}$ in the chronic RA is presented below.

1. *Chronic Assessment.* Is $PEC_{sw;max}$ (of highest available tier) > $RAC_{sw;ch}$ (of highest available tier)?

   Yes: Go to 2
   No: Low chronic risk

2. Is the $RAC_{sw;ch}$ derived from a test with algae or from a long-term (≥ 7 days) test with another water organism and the following conditions apply (i) loss of the a.s. from water is more than 20 % of nominal at the end of the exposure period and (ii) the toxicity estimate (e.g. EC_{10} or NOEC) is expressed in terms of nominal/initially measured concentration of the a.s.?

Yes: $PEC_{sw;twa}$ not appropriate (low risk not demonstrated)
No: Go to 3

3. Is the $RAC_{sw;ch}$ based on treatment-related responses of the relevant test species early in the chronic test (e.g. during the initial 96 h observed mortality/immobility in tests with animals, or 50 % reduction in growth rate in tests with macrophytes, in the treatment level above the one from which the $RAC_{sw;ch}$ is derived) or is the acute to chronic ratio (acute $L(E)C_{50}$/chronic NOEC or acute $L(E)C_{50}$/chronic EC_{10}) based on immobility or mortality < 10?

   Yes: $PEC_{sw;twa}$ not appropriate (low risk not demonstrated)
   No: Go to 4

4. Is it demonstrated by the applicant that for the organisms and the PPP under evaluation and/or a PPP with a similar toxic mode of action (read-across information), the following phenomena are not likely: (i) latency of effects due to short-term exposure and (ii) the co-occurrence of exposure and specific sensitive life stages that last only a short time.

   Yes: Go to 5
   No: $PEC_{sw;twa}$ not appropriate (low risk not demonstrated)

5. Is $PEC_{sw;7d\text{-}twa}$ (of highest available tier) > $RAC_{sw;ch}$ (of highest available tier)?

   Yes: Go to 6
   No: Low risk demonstrated

6. Are experimental (or TK/TD modelling when guidance is available) data available that demonstrate that, for the species, a larger time window for the $PEC_{sw;twa}$ may be used (not exceeding the duration of the tier 1 chronic test that triggered the risk)?

   Yes: Go to 4 and replace the $PEC_{sw;7d\text{-}twa}$ by another appropriate $PEC_{sw;twa}$
   No: Low risk not demonstrated

For invertebrates, fish and macrophytes, a default 7-day time window is proposed for the $PEC_{sw,twa}$ if the TWA approach is deemed acceptable. It may be justified to lengthen or shorten the default 7-day TWA period of the PEC if justified with appropriate scientific data (for further details see section 4.5.1).

In the sections below, a concise summary will be presented of effect tiers 1 to 3. However, before starting the acute and chronic effect assessment for RAC derivation it is essential to gather information on the predicted exposure profiles for the pesticide of concern in the relevant edge-of-field water bodies on the basis of FOCUS surface water (see chapter 6) and/or Member State-specific exposure scenarios. This is of particular importance when evaluating refined exposure tests (tier 2C) and micro-/mesocosm experiments (tier 3), since these studies are able to address the effects of the predicted field exposure profile in a more realistic way. The rationale for this is presented in section 7.1.

2.1.3. Tier 1 RAC_{sw} derivation on the basis of standard test species

The specific data requirements for a.s. and PPPs under Regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market are laid down in Commission Regulation (EU) 283/2013[6] for the dossier to be submitted for the approval of a.s. contained in PPPs and in Commission Regulation (EU) 284/2013[7] for the authorisation of PPPs. The obligatory toxicity tests that should be provided for pesticides are presented below in Table 1 to Table 3. A more detailed

[6] Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, p. 1–84.

[7] Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, p. 85–152.

description of the tier 1 data requirements is presented in chapter 7. Note that in the tables below, the tests with algae and macrophytes are placed under the chronic RA since these tests comprise the complete life cycle, or a large part of the life cycle, of these organisms, although the toxicity endpoint selected is the EC_{50}.

Table 1: The obligatory toxicity tests that should be provided for pesticides with an insecticidal mode of action

	Standard test species	Duration	Endpoint	RAC
Acute effect assessment	*Daphnia* sp. (*D. magna* preferred)	48 h	EC_{50}	EC_{50}/100
	Additional arthropod [(a)] (e.g. *Chironomus* sp. or *Americamysis bahia*)	48 h	EC_{50}	EC_{50}/100
	Oncorhynchus mykiss	96 h	LC_{50}	LC_{50}/100
Chronic effect assessment	*Daphnia* sp. or additional arthropod [(b)]	21 d	EC_{10} (NOEC)	EC_{10}/10
	Chironomus spp. [(c)]	20–28 d	EC_{10} (NOEC)	EC_{10}/10
	Early life stage test or full life cycle test with fish [(d)]	Variable	EC_{10} (NOEC)	EC_{10}/10
	Green alga (e.g. *Pseudokirchneriella subcapitata*)	72 h[(e)]	E_rC_{50}[(f)]	E_rC_{50}/10

(a): The PPR Panel recommends to preferably use a *Chironomus* test, if data on *A. bahia* are not already available.

(b): Preferably the most sensitive standard test arthropod (*Daphnia*, *Chironomus*, *Americamysis*) from the acute tier 1 dataset should be selected as test species in the chronic effect assessment. If in the acute assessment a certain standard test arthropod is a factor of 10 more sensitive, the PPR Panel advices to always perform a chronic test with this arthropod.

(c): Obligatory only if the substance partitions to sediment and/or when the substance interferes with moulting hormones (e.g. insect growth regulators). The substance is considered to partition to the sediment if the water/sediment study shows > 10 % of applied radioactivity at or after day 14 present in the sediment and the chronic *Daphnia* test (or other comparable study with e.g. *Chironomus*) shows a EC_{10}/NOEC of < 0.1 mg/L.

(d): Early life stage test required where exposure of surface water is possible and the substance does not hydrolyse instantly ($DegT_{90}$> 1 d), unless a fish full life cycle (FFLC) test is provided. An FFLC test may be required depending upon the persistence and bioaccumulative potential of the substance. The Panel recommends that FFLC tests may be required where the BCF is > 1 000, the elimination during the 14 day depuration phase in the bioconcentration study is < 95 % or the substance is stable in water or sediment ($DegT_{90}$ > 100 days). Long-term exposure may also occur for substances which show degradation in water and sediment if leaching from drainpipes contributes significantly to the exposure in surface water. So if long-term exposure is expected based on the predicted field exposure profile, a FFLC study might be required as well. However, it is not yet possible to provide rules of thumb for the significance of leaching from drainpipes based on the $DegT_{50}$ in soil, the K_{om} and other relevant substance characteristics. Development of such rules of thumb may be helpful for the RA. A FFLC will be required if there are indications that the substance has endocrine-mediated effects in a fish screening test.

(e): The test duration based on other Technical Guidelines (i.e. EPA OPPTS 850.4500) for algae is 96 h instead of 72 h. Endpoints from these tests are also acceptable for deriving a chronic RAC for algae.

(f): Growth rate (r) is the preferred endpoint. Other, usually more sensitive endpoints such as yield may also be used if growth rate endpoints are not provided.

Table 2: The obligatory toxicity tests that should be provided for pesticides with a herbicidal modeof action

	Standard test species	Duration	Endpoint	RAC
Acute effect assessment	*Daphnia* sp. (*D. magna* preferred)	48 h	EC_{50}	EC_{50}/100
	Oncorhynchus mykiss	96 h	LC_{50}	LC_{50}/100
Chronic effect assessment	Green alga (e.g. *Pseudokirchneriella subcapitata*)	72 h [(d)]	E_rC_{50} [(e)]	E_rC_{50}/10
	Additional non-green alga (e.g. diatom *Navicula pelliculosa*)	72 h [(d)]	E_rC_{50} [(e)]	E_rC_{50}/10
	Lemna sp or *Myriophyllum* sp. or *Glyceria maxima* [(a)]	7–14 d	E_rC_{50} [(e)]	EC_{50}/10
	Daphnia sp.	21 d	EC_{10} (NOEC)	EC_{10}/10

	Early life stage test or full life cycle test with fish [(b)]	–	EC_{10} (NOEC)	$EC_{10}/10$
	Chironomus sp. or *Lumbriculus* sp. [(c)]	20–28 d	EC_{10} (NOEC)	$EC_{10}/10$

(a): *Lemna* sp. is the default macrophyte test species. In case *Lemna* and algae are apparently not sensitive to the herbicidal product (e.g. EC_{50} > 1mg/L), or if the herbicide simulates a plant growth hormone, a rooted macrophyte is required (preferably *Myriophyllum*). It is advised to test *Glyceria* in the case of a herbicide that primarily affects monocots in terrestrial plant trials.

(b): Early-life stage test required where exposure of surface water is possible and the substance does not hydrolyse instantly ($DegT_{90}$> 1 d), unless a fish full life-cycle (FFLC) test is provided. An FFLC test may be required depending upon the persistence and bioaccumulative potential of the substance. The PPR Panel recommends that FFLC-tests may be required where the BCF is > 1 000, the elimination during the 14-day depuration phase in the bioconcentration study is < 95 % or the substance is stable in water or sediment ($DegT_{90}$ > 100 days). Long-term exposure may also occur for substances which show degradation in water and sediment if leaching from drainpipes contributes significantly to the exposure in surface water. Thus, if long-term exposure is expected based on the predicted field exposure profile, a FFLC study might be required as well. However, it is not yet possible to provide rules of thumb for the significance of leaching from drainpipes based on the $DegT_{50}$ in soil, the K_{om} and other relevant substance characteristics. Development of such rules of thumb may be helpful for the RA. A FFLC will be required if there are indications that the substance has endocrine-mediated effects in a fish screening test.

(c): Substance accumulates in sediment. The substance is considered to accumulate in the sediment if the water/sediment study shows > 10 % of applied radioactivity at or after day 14 present in the sediment and the chronic *Daphnia* test (or other comparable study with e.g. *Chironomus*) shows a EC_{10}/NOEC of < 0.1 mg/L.

(d): The test duration based on other Technical Guidelines (i.e. EPA OPPTS 850.4500) for algae is 96 h instead of 72 h. Endpoints from these tests are also acceptable for deriving a chronic RAC for algae.

(e): Growth rate (r) is the preferred endpoint. Other, usually more sensitive, endpoints such as yield may also be used if growth rate endpoints are not provided.

Table 3: The obligatory toxicity tests that should be provided for other pesticides.

	Standard test species	**Duration**	**Endpoint**	**RAC**
Acute effect assessment	*Daphnia* sp. (*D. magna* preferred)	48 h	EC_{50}	$EC_{50}/100$
	Oncorhynchus mykiss	96 h	LC_{50}	$LC_{50}/100$
Chronic effect assessment	Green alga (e.g. *Pseudokirchneriella subcapitata*)	72 h [(c)]	E_rC_{50} [(d)]	$E_rC_{50}/10$
	Daphnia sp.	21 d	EC_{10} (NOEC)	$EC_{10}/10$
	Early life stage test or full life cycle test with fish [(a)]	–	EC_{10} (NOEC)	$EC_{10}/10$
	Chironomus sp. or *Lumbriculus* sp. [(b)]	20–28 d	EC_{10} (NOEC)	$EC_{10}/10$

(a): Early life-stage test required where exposure of surface water is possible and the substance does not hydrolyse instantly ($DegT_{90}$> 1 d), unless a fish full life-cycle (FFLC) test is provided. An FFLC test may be required depending upon the persistence and bioaccumulative potential of the substance. The Panel recommends that FFLC-tests may be required where the BCF is > 1 000, the elimination during the 14 day depuration phase in the bioconcentration study is < 95 % or the substance is stable in water or sediment ($DegT_{90}$ > 100 days). Long-term exposure may also occur for substances which show degradation in water and sediment if leaching from drainpipes contributes significantly to the exposure in surface water. Thus, if long-term exposure is expected based on the predicted field exposure profile, a FFLC study might be required as well. However, it is not yet possible to provide rules of thumb for the significance of leaching from drainpipes based on the $DegT_{50}$ in soil, the K_{om} and other relevant substance characteristics. Development of such rules of thumb may be helpful for the RA. A fish FFLC will be required if there are indications that the substance has endocrine-mediated effects in a fish screening test.

(b): Substance accumulates in sediment. The substance is considered to accumulate in the sediment if the water/sediment study shows > 10 % of applied radioactivity at or after day 14 present in the sediment and the chronic *Daphnia* test (or other comparable study with e.g. *Chironomus*) shows a EC_{10}/NOEC of < 0.1 mg/L. For fungicides the PPR Panel recommends the *Lumbriculus* test.

(c): The test duration based on other Technical Guidelines (i.e. EPA OPPTS 850.4500) for algae is 96 h instead of 72 h. Endpoints from these tests are also acceptable for deriving a chronic RAC for algae.

(d): Growth rate (r) is the preferred endpoint. Other, usually more sensitive, endpoints such as yield and biomass may also be used if growth rate endpoints are not provided.

2.1.4. Tier 2 RAC_{sw} derivation on the basis of additional laboratory toxicity tests

If, besides the basic data requirements presented in Table 1 to Table 3, additional laboratory toxicity tests are provided, a tier 2 effect assessment may be performed. In the GD, three different tier 2 effect assessments are described:

- Tier 2A: The Geomean assessment factor (AF) approach (see section 8.3 for a detailed description);
- Tier 2B: The species sensitivity distribution (SSD) approach (see section 8.3);
- Tier 2C: The refined exposure laboratory test AF approach (see section 9.2).

2.1.4.1. Tier 2A: The Geomean-AF approach

The PPR Panel of EFSA advices to apply the Geomean-AF approach as summarised in Table 4. In this approach, the Geomean $L(E)C_{50}$ or Geomean $NOEC/EC_{10}$ values for species belonging to the same taxonomic group (e.g. separately for crustaceans, insects, fish, green algae, diatoms, monocotyledonous macrophytes, dicotyledonous macophytes) are calculated and the AF of the tier 1 effect assessment is applied.

Table 4: Proposal for the derivation of RACs for aquatic organisms when a limited number of additional single species toxicity tests is available. When applying this approach scientific arguments should be given why the selected toxicity data (on which the Geomean is based) concern the same taxonomic group relevant for the RA. If more data than indicated in the second column (≥ 5 or ≥ 8) are available, the Geomean approach could still be applied, but it is recommended to preferably apply the SSD approach (see section 8.3).

Taxonomic group	Number of toxicity data for different taxa of the relevant taxonomic group	RAC (Geomean-EC_x/AF)	Field exposure concentration (PEC)
Acute RA			
Aquatic vertebrates[(a)]	< 5 acute LC_{50s}	Geomean $LC_{50}/100$ [(d)]	$PEC_{sw;max}$
Invertebrates [(b)]	< 8 acute EC_{50s}	Geomean $EC_{50}/100$ [(d)]	$PEC_{sw;max}$
Chronic RA			
Aquatic vertebrates[(a)]	< 5 chronic EC_{10s} (or chronic NOECs)	Geomean $EC_{10}/10$ [(d, e)]	$PEC_{sw;max}$ or $PEC_{sw;twa}$
Invertebrates [(b)]	< 8 chronic EC_{10s} (or chronic NOECs)	Geomean $EC_{10}/10$ [(d, e)]	$PEC_{sw;max}$ or $PEC_{sw;twa}$
Primary producers[(c)]	< 8 EC_{50s}	Geomean $EC_{50}/10$ [(d, e)]	$PEC_{sw;max}$

(a): i.e. fish or amphibians.
(b): i.e. separately for crustaceans and insects in the case of insecticides, and a more specific taxonomic group in the case of fungicides, unless it is demonstrated that certain taxonomic groups can be combined.
(c): i.e. separately for green algae, diatoms, blue-green algae, monocotyledonous macrophytes and dicotyledonous macrophytes in the case of herbicides or fungicides with a herbicidal mode of action, unless it is demonstrated that certain taxonomic groups can be combined. E_rC_{50s} on the basis of growth rate and the most sensitive ecologically relevant endpoint are preferred (in accordance with the relevant OECD guidelines). Yield endpoints may also be used if growth rate endpoints are not provided.
(d): Of the different taxonomic groups, the lowest Geomean value is selected (e.g. the lowest value for insects or crustaceans in the case of insecticides; the lowest value for green algae, diatoms, blue-green algae or macrophytes in the case of herbicides)
(e): When applying the Geomean approach to chronic toxicity data, comparable endpoints should be used within the same taxonomic group.

There is a possibility that the outcome of the geometric mean approach is biased by introducing insensitive species. In the case of differences in sensitivity of 1 or 2 orders of magnitude (factor 10–100) an assessment of this possibility has to be made. If the most sensitive species is more than a

factor of 10 (for plants and chronic tests) or 100 (for acute invertebrate and fish tests) below the geometric mean of all the tested species within the relevant taxonomic group, a weight of evidence approach should be applied or more toxicity data may be generated. In addition, if on the basis of the toxic mode of action of the a.s. (e.g. insect growth regulators and acute tests with insects) delayed effects can be expected that are not covered by the standard duration of the acute toxicity test, the Geomean cannot be used or should be based on prolonged toxicity tests.

2.1.4.2. Tier 2B: The Species Sensitivity Distribution (SSD) approach

The PPR Panel of EFSA advices to apply the SSD approach as summarised in Table 5 for aquatic invertebrates and primary producers, and in Table 6 for aquatic vertebrates (see section 8.4). In this approach, the median HC_5 (hazardous concentration to 5 % of the tested species that is predicted with 50 % certainty) and the lower limit HC_5 values ($LLHC_5$: hazardous concentration to 5 % of the tested species that is predicted with 95 % certainty) are derived from the SSD curves that are constructed with at least 8 representative toxicity data for different non-vertebrate species or with at least 5 representative toxicity data for different fish and/or amphibian species. For $RAC_{sw;ac}$ derivation, acute toxicity data of the relevant sensitive taxonomic groups should be used to construct the SSD. See section 8.4.3 for further information on species selection to construct SSDs for insecticides, herbicides and fungicides. If the specific toxic mode-of-action of the compound is likely to result in latent effects (e.g. as demonstrated for some insect growth regulators) the SSD should be constructed with EC_{50} values derived from prolonged acute toxicity tests. In prolonged acute toxicity tests, the observation of treatment-related responses is continued after the test organisms are transferred to a clean medium. Chronic toxicity data of the relevant sensitive taxonomic groups should be used to construct the SSD for $RAC_{sw;ch}$ derivation. In Table 5 and Table 6, a distinction is made between RAC derivation for vertebrates and non-vertebrates on the basis of the SSD approach, since a higher protection level is required in the acute effect assessment for fish and amphibians (avoidance of mortality and suffering due to acute toxicity of pesticide exposure).

For primary producers, the PPR Panel recommends to calculate the SSD-RAC both on the basis of the median HC_5 and the application of an AF of 3.

For invertebrates, the PPR Panel recommends to calculate the SSD-RAC based on acute effect data both on the basis of the median HC_5 and the application of an AF of 3 to 6. Note that the size of the AF is based on calibration of median HC_5 estimates with threshold concentrations observed in micro-/mesocosms (see 8.4.4), the extrapolation between threshold concentrations observed in micro-/mesocosms and the field (see 9.3.5), and the size of the tier 1 AF (i.e. in the acute tier 1 assessment for invertebrates, an AF of 100 is used, while in the tier 1 effect assessment for plants, an AF of 10 is used). The following aspects may be further considered in selecting the size of the AF to derive an SSD-$RAC_{sw;ac}$ for invertebrates (within the ranges proposed).

1. *The quality of the acute toxicity data used to construct the SSD*. If the toxicity data comprise several different genera/families/orders of the potentially sensitive taxonomic group (see section 8.4.3 for further guidance), including Ephemeroptera/Plecoptera/Trichoptera taxa (EPT) for insecticides, a lower AF in the proposed range may be selected. However, if another valid SSD can be constructed with a more limited dataset containing the most sensitive species, and the HC_5 derived from this SSD curve is lower than that of the SSD curve using toxicity data for a wider array of taxa, a higher AF in the proposed range may be selected to be applied to the SSD from the wider set.

2. *The lower limit value of the HC_5*. If the lower limit HC_5 derived from the curve is less than 1/3 of the median HC_5, a higher AF in the proposed range may be warranted.

3. *The lower tier RACs on the basis of standard toxicity data (tier 1), standard and additional toxicity data (Geomean approach) and tier 3 data*. The size of the AF should ideally not result in an SSD-$RAC_{sw;ac}$ higher than the tier 3 RAC derived from effect class 1 and 2 of micro-

/mesocosm studies, nor should it result in an SSD-$RAC_{sw;ac}$ lower than the tier 1 $RAC_{sw;ac}$ on the basis of standard test species and/or the Geomean-$RAC_{sw;ac}$ and/or method 3 to 5 (EFSA, 2006a) on the basis of the same toxicity data that were used to construct the SSD. Note that according to EFSA (2006a), the Geomean approach aims to achieve the same average level of protection as in the tier 1 effect assessment but can be predicted more accurately because of the availability of additional toxicity data for the relevant taxonomic groups.

4. *The position of the toxicity data in the lower tail of the SSD (around the HC_5)*. If in the lower tail the toxicity data, overall, are positioned on the right side of the SSD curve, the derived HC_5 estimate may be considered relatively 'conservative' for the most sensitive species. This may be a reason to adopt a lower AF from the proposed range. In contrast, if in the lower tail the toxicity data are, overall, positioned on the left side of the SSD curve, this may be a reason to adopt a higher AF from the proposed range.

5. *The steepness of the SSD curve*. In the case of a relatively steep SSD curve (e.g. less than a factor of 100 between lowest and highest $L(E)C_{50}$ value used to construct the SSD curve), a higher AF from the proposed range is recommended since exposure concentrations that exceed the $RAC_{sw;ac}$ may have ecotoxicological consequences for a larger number of taxa.

6. *Read-across information for compounds with a similar toxic mode of action*. For a PPP with a well-known mode of action, sufficient higher tier information on related compounds (e.g. organophosphates) may be available that allows the evaluation of the predictive value of the median HC_5 and/or lower limit of the HC_5 for possible effects in micro-/mesocosms. This information may be used to select an appropriate AF within the proposed range.

7. *Considering information on chronic effects*. If the acute to chronic ratio (acute EC_{50}/chronic EC_{10}) is larger than 10, then an AF in the higher range may be warranted.

Table 5: Proposal for the derivation of a RAC in edge-of-field surface waters, based on hazardous concentrations derived from species sensitivity distributions with aquatic invertebrates and/or plants.

Type of effect/risk assessment	Relevant PEC	Hazardous concentration	AF to derive RAC from hazardous concentration
Acute and chronic effect/risk assessment for invertebrates and single and repeated pulse exposure	$PEC_{sw;max}$	***Latency of effects not expected***[(a)] Median acute HC_5 (based on acute LC_{50} or EC_{50} data)[(b)]	3–6
		Latency of effects expected *(e.g. insect growth regulators).* Median acute HC_5 (based on acute LC_{50} or EC_{50} data from prolonged acute toxicity tests[(c)]) *or* precautionary approach instead of the 2 options above: apply chronic SSD (see below)	3–6
Chronic effect/risk assessment for invertebrates and long-term exposure (concentrations during relevant time window	$PEC_{sw;max}$ or $PEC_{sw;twa}$	Median chronic HC_5 (based on chronic NOEC and/or EC_{10} data)	3

Type of effect/risk assessment	Relevant PEC	Hazardous concentration	AF to derive RAC from hazardous concentration
> 10 % of initial peak concentration)			
Effect/risk assessment for primary producers	$PEC_{sw;max}$ or, in the case of macrophyte SSDs only, $PEC_{sw;twa}$	Median HC_5 (based on EC_{50} data)[(b,d)]	3

(a): This has to be demonstrated by the applicant, see further section 4.5.1. For example, by read-across for substances with similar toxic mode of action, prolonged acute toxicity tests, and information from micro-/mesocosm studies for similar compounds with a longer-term observation period after exposure.
(b): For types of PPPs evaluated by Maltby et al. (2005; 2009) and Van den Brink et al. (2006).
(c): In prolonged acute toxicity tests the observation of treatment-related responses is continued after the test organisms are transferred to a clean medium.
(d): E_rC50s on the basis of growth rate and the most sensitive ecologically relevant endpoint are preferred (in accordance with the relevant OECD guidelines). Yield endpoints may also be used if growth rate endpoints are not provided.

Table 6: Proposal for the derivation of an RAC for edge-of-field surface waters, based on hazardous concentrations derived from species sensitivity distributions with fish (and other aquatic vertebrates). Information of possible latency of effect may be obtained on the basis of knowledge on the specific toxic mode of action and, read-across information. To avoid unnecessary testing with aquatic vertebrates for animal welfare considerations, the conduct of prolonged acute toxicity tests to demonstrate latency is not considered

Type of effect/risk assessment	Relevant PEC	Hazardous concentration	AF to derive RAC from hazardous concentration
Acute effect/risk assessment	$PEC_{sw;max}$	***Latency of effects not expected***[(a)] Median acute HC_5 (based on 96 h NOEC and/or acute LC_{10} data) *or* Median acute HC_5 (based on 96 h LC_{50} or EC_{50} data) *or* ***If latency of effects is expected*** go to chronic effect assessment (see below)	3 9
Chronic effect/risk assessment	$PEC_{sw;max}$ or $PEC_{sw;twa}$	Median chronic HC_5 (based on chronic NOEC and/or EC_{10} data)	3

(a): This has to be demonstrated by the applicant, see further section 4.5.1. For example, by read-across for substances with similar toxic mode of action, prolonged acute toxicity tests for similar compounds with a longer-term observation period after exposure.

2.1.5. Tier 2C: refined exposure laboratory test-AF approach

The PPR Panel of EFSA proposes to explore a higher tier RAC derivation on the basis of the refined exposure laboratory-AF approach if predicted (modelled) exposure profiles for edge-of-field surface waters differ considerably from exposure regimes in standard toxicity studies and if the $PEC_{sw;twa}$ cannot be used in the chronic RA (for a more detailed description see section 9.2). Refined exposure

laboratory tests are usually performed with the tier 1 standard test species that drive the aquatic risks, and are designed in such a way that the exposure in these tests more realistically resembles the field exposure conditions. Nevertheless, in order to consider them as an appropriate higher tier effect assessment approach, the refined exposure tests should simulate a realistic worst-case exposure relative to that predicted for the edge-of-field, and they should be long enough to allow the expression of the maximum effects. In acute RAs, this usually requires prolonged acute refined exposure toxicity tests. In chronic RAs, the duration of the refined exposure tests is usually similar to that of the chronic tier 1 standard toxicity test, but may be longer (e.g. in the case of tests with algae). Note that the RACs derived from refined exposure toxicity tests should always be expressed in terms of peak exposure concentration in these tests, and that these RACs should always be compared with the $PEC_{sw;max}$.

A summary of the $RAC_{sw;ac}$ and $RAC_{sw;ch}$ derivation on the basis of refined exposure laboratory tests, and their use in the RA, is presented in Table 7.

Table 7: Derivation of RACs in edge-of-field surface waters, based on refined exposure laboratory toxicity tests with standard test species

Type of effect/risk assessment	Relevant PEC	Endpoint of refined exposure toxicity test with standard test species expressed in terms of peak exposure concentration in test system	RAC
Acute effect/risk assessment	$PEC_{sw;max}$	$L(E)C_{50}$ (animal tests)	$L(E)C_{50}/100$
Chronic effect/risk assessment	$PEC_{sw;max}$	EC_{50} (plant tests)	$EC_{50}/10$
		EC_{10}/NOEC (animal tests)	$EC_{10}/10$

2.1.6. Tier 3 RAC_{sw} derivation on the basis of micro-/mesocosm tests

The requirements for the conduct and interpretation of micro-/mesocosm tests for RAC_{sw} derivation are described in detail in section 9.3. To evaluate the scientific reliability of micro-/mesocosm experiments, the following questions should be addressed:

1. *Is the test system adequate and does the test system represent a realistic freshwater community?* (Trophic levels; taxa richness and abundance of (key and sensitive) species; representativeness of the biological traits with respect to vulnerability).

2. *Is the description of the experimental set-up adequate and unambiguous?* (ANOVA or regression design; overall characterisation of the experimental ecosystem/community simulated; measurement endpoints; sampling frequency; sampling techniques).

3. *Is the exposure regime adequately described?* (Method of application of the test substance; relevance for predicted exposure profile in the field; concentration in the application solution; dynamics in exposure concentrations in relevant compartments (e.g. water, sediment); detection limits).

4. *Are the investigated endpoints sensitive and in accordance with the working mechanisms of the compound, and with the results of the first tier studies?* (Compare selected measurement endpoints with the species potentially at risk as indicated by the lower tiers.)

5. *Is it possible to evaluate the observed effects statistically and ecologically?* (Univariate and multivariate techniques applied; unambiguous concentration–response relationships; statistical power of the test; ecological relevance of the statistical output).

Furthermore, the criteria for RAC derivation from micro-/mesocosm tests on the basis of the ETO and the ERO are presented below in Decision scheme C. In Decision scheme C, reference is made to effect

class concentrations for the most sensitive measurement endpoints derived from micro-/mesocosm tests. The following effect classes are important in the derivation of ETO-RAC_{sw} and ERO-RAC_{sw} values.

- Effect class 1 (no treatment-related effects demonstrated for the most sensitive endpoints). No (statistically and/or ecologically significant) effects observed as a result of the treatment. Observed differences between treatment and controls show no clear causal relationship.

- Effect class 2 (slight effects). Effects concern short-term and quantitatively restricted responses usually observed at individual samplings only.

- Effect class 3A (pronounced short-term effects (< 8 weeks), followed by recovery). Clear response of endpoint, but full recovery of affected endpoint within 8 weeks after the 1st application or, in the case of delayed responses and repeated applications, the duration of the effect period is less than 8 weeks and followed by full recovery. Effects observed at some subsequent sampling instances.

A more elaborate description of all effect classes is given in section 9.3.3.1.

The PPR Panel of EFSA proposes the procedure presented in Table 8 and Table 9 to derive an ETO-RAC_{sw} on the basis of micro-/mesocosm tests, while that for ERO-RAC_{sw} derivation is presented in

Table **10** and Table 11.

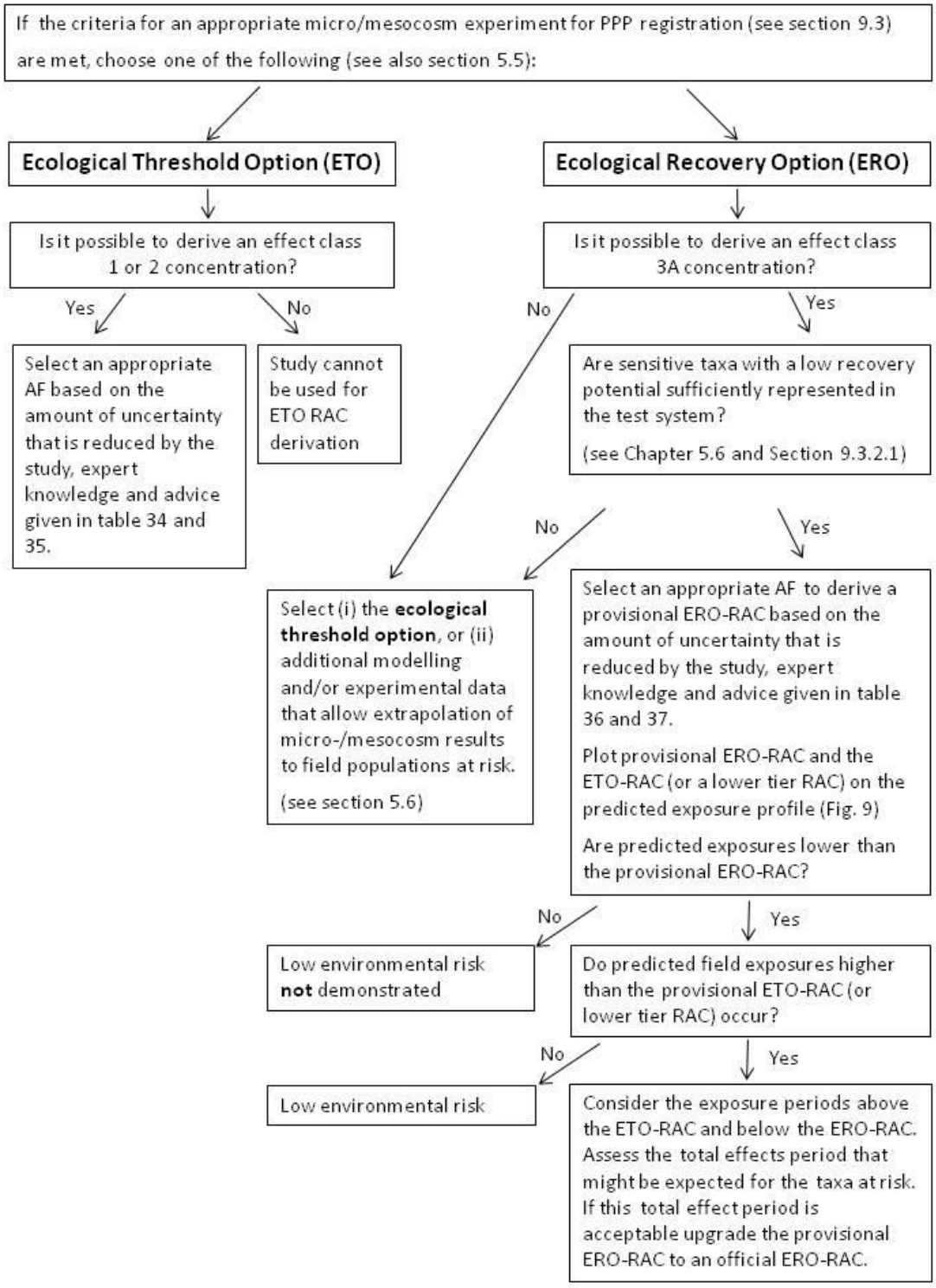
The PPR Panel of EFSA proposes the procedure presented in

Table **10** and Table 11 to derive an ERO-RAC_{sw} on the basis of micro-/mesocosm tests.

Wherever possible, both an ETO-RAC_{sw} and an ERO-RAC_{sw} should be derived from the same micro-/mesocosm study.

A range in size of AF is proposed to derive the ETO-$RAC_{sw;ac}$ (Table 8), the ETO-$RAC_{sw;ch}$ (Table 9), the ERO-$RAC_{sw;ac}$ (

Table **10**) and the ERO-$RAC_{sw;ch}$ (Table 11). Guidance in selecting the size of the AF is provided in section 9.3.5.1.



Decision scheme C: Summary flow chart for the derivation of RACs from appropriate micro-/mesocosm experiments on the basis of the ecological threshold option (ETO-RAC) or the ecological recovery option (ERO-RAC)

Table 8: Proposal for the derivation of the $RAC_{sw;ac}$ (triggered by tier 1 acute core data) addressing the ETO on the basis of an appropriate micro-/mesocosm experiment. Note that, in the same study, several treatment levels may result in effect class 1 responses for sensitive measurement endpoints. In that case, the highest treatment level showing an overall effect class 1 response should be selected for ETO-RAC derivation. Alternatively, if, in the same study, several treatments result in effect class 2 responses in the first instance, the lowest treatment level showing an overall effect class 2 response should be selected for ETO-RAC derivation. On a case-by-case basis, and with expert judgement, it may be decided to select a higher treatment level as overall effect class 2 concentration.

	Assessment factor for ETO-$RAC_{sw;ac}$ derivation (ecological threshold option)	**Field exposure concentration to compare with the $RAC_{sw;ac}$**
Effect class 1 concentration *Is rate of dissipation of the a.s. in test system realistic to worst case when compared with that predicted for the field?* *Yes*: Base effect estimate on nominal or measured peak concentration in test system. *No*: Base effect estimate on, for example, the initial 48 h TWA concentration in test system or apply appropriate extrapolation techniques.	2[(a)]	$PEC_{sw;max}$
Effect class 2 concentration *Is rate of dissipation of the a.s. in test system realistic to worst case when compared with that predicted for the field?* *Yes*: Base effect estimate on nominal or measured peak concentration in test system. *No*: Base effect estimate on, for example, the initial 48 h TWA concentration in test system or apply appropriate extrapolation techniques.	2–3[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available, the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. If the available micro-/mesocosm studies are of the same quality, the AF may be applied to the geometric mean value of the effect class 1 or effect class 2 concentrations derived from the different studies. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

Table 9: Proposal for the derivation of the $RAC_{sw;ch}$ (triggered by tier 1 chronic core data) addressing the ETO on the basis of an appropriate micro-/mesocosm experiment. Note that, in the same study, several treatment levels may result in effect class 1 responses for sensitive measurement endpoints. In that case the highest treatment level showing an overall effect class 1 response should be selected for ETO-RAC derivation. Alternatively, if, in the same study, several treatments result in effect class 2 responses in first instance the lowest treatment level showing an overall effect class 2 response should be selected for ETO-RAC derivation. On a case-by-case basis, and with expert judgement, it may be decided to select a higher treatment level as overall effect class 2 concentration

	Assessment factor for ETO-$RAC_{sw;ch}$ derivation (ecological threshold option)	**Field exposure concentration to compare with the $RAC_{sw;ch}$**
Effect class 1 concentration Based on time weighted average concentration in test system during the application period	2[(a)]	$PEC_{sw;max}$ or $PEC_{sw;twa}$. Based on expert judgement by considering the criteria mentioned

	Assessment factor for ETO-$RAC_{sw;ch}$ derivation (ecological threshold option)	**Field exposure concentration to compare with the $RAC_{sw;ch}$**
		in chapter 4
Based on nominal or peak concentration in test system if the long-term exposure regime (e.g. owing to repeated pulses) is realistic to worst case compared with the predicted field exposure profile	2[(a)]	$PEC_{sw;max}$
Effect class 2 concentration		
Based on time weighted average concentration in test system during the application period	2–3[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$ or $PEC_{sw;twa}$. Based on expert judgement by considering the criteria mentioned in chapter 4
Based on nominal or peak concentration in test system if the long-term exposure regime (e.g. owing to repeated pulses) is realistic to worst case compared with the predicted field exposure profile	2–3[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available, the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. If the available micro-/mesocosm studies are of the same quality, the AF may be applied to the geometric mean value of the effect class 1 or effect class 2 concentrations derived from the different studies. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

Table 10: Proposal for the derivation of the $RAC_{sw;ac}$ (triggered by tier 1 acute core data) addressing the ERO on the basis of an appropriate micro-/mesocosm experiment. If, in the same study, several treatments result in effect class 3A responses for sensitive measurement endpoints in the first instance, the lowest treatment level showing an overall effect class 3A response should be selected to derive the ERO-RAC. On a case-by-case basis, and with expert judgement, it may be decided to select a higher treatment level as overall effect class 3A concentration

	Assessment factor for ERO-$RAC_{sw;ac}$ derivation (ecological recovery option)	**Field exposure concentration to compare with the $RAC_{sw;ac}$**
Effect class 3A concentration		
Maximum magnitude of temporal effects may be medium to large. *Is rate of dissipation of the a.s. in test system realistic to worst case when compared with that predicted for the field?* *Yes*: Base effect estimate on nominal or measured peak concentration in test system. *No*: Base effect estimate on, for example, the	3–4[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

	Assessment factor for ERO-$RAC_{sw;ac}$ derivation (ecological recovery option)	**Field exposure concentration to compare with the $RAC_{sw;ac}$**
initial 48 h TWA concentration in test system or, apply appropriate extrapolation techniques or, consider the ecological threshold option (Table 8).		

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available, the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. If the available micro-/mesocosm studies are of the same quality, the AF may be applied to the geometric mean value of the effect class 3A concentrations derived from the different studies. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

Table 11: Proposal for the derivation of the $RAC_{sw;ch}$ (triggered by tier 1 chronic core data) addressing the ERO on the basis of an appropriate micro-/mesocosm experiment. If, in the same study, several treatments result in effect class 3A responses for sensitive measurement endpoints in the first instance, the lowest treatment level showing an overall effect class 3A response should be selected to derive the ERO-RAC. On a case-by-case basis, and with expert judgement, it may be decided to select a higher treatment level as overall effect class 3A concentration

	Assessment factor for ERO-$RAC_{sw;ch}$ derivation (ecological recovery option)	**Field exposure concentration to compare with the $RAC_{sw,ch}$**
Effect class 3A concentration Based on time weighted average concentration in test system during the application period	3–4.[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$ or $PEC_{sw;twa}$. Based on expert judgement by considering the criteria mentioned in chapter 4
Based on nominal or peak concentration in test system if the long-term exposure regime (e.g. owing to repeated pulses) is realistic to worst case compared with the predicted field exposure profile	3–4.[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available, the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. If the available micro-/mesocosm studies are of the same quality, the AF may be applied to the geometric mean value of the effect class 3A concentrations derived from the different studies. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

2.2. Bioconcentration and secondary poisoning

Some compounds in the water have the tendency to accumulate in the tissue of fish or in the tissue of other organisms. This tendency of a compound is often expressed in a bioconcentration factor (BCF). The equilibrium concentration for a compound in fish can be estimated by multiplying the concentration of the compounds in the surrounding water by the fish BCF for that particular

compound. At long exposure times (equilibrium), the BCF also equals the ratio of the uptake rate constant and depuration and elimination rate constant (Mackay, 1982).

Bioaccumulation often correlates with lipophilicity. For organic chemicals, a log K_{ow} of ≥ 3 indicates a potential for bioaccumulation. The stability of a compound is another indicator for accumulation. The compound is considered stable when less than 90 % loss of the original substance over 24 h via hydrolysis has been noted (see section 7.6).

The regulatory acceptable concentrations for secondary poisoning (RAC_{sp}) for birds and mammals eating fish out of the surface water contaminated with a PPP can be assessed in the following way (see section 7.6.3):

$$RAC_{SP} = \frac{NOAEL_{bird}}{5 \times 0.159 \times BCF_{fish}} \text{ or } \frac{NOAEL_{mammal}}{5 \times 0.138 \times BCF_{fish}}$$

where:

RAC_{sp}: regulatory acceptable concentration in water for secondary poisoning (mg/L);

NOAEL: relevant long-term no-adverse-effect-level for birds or mammals (mg/kg body weight per day);

BCF_{fish}: whole body bioconcentration factor in fish (L/kg);

5 is the AF; 0.159 and 0.138 are multiplication factors based on a 1 000-g bird eating 159 g of fish per day and a 3 000-g mammal eating 415 g of fish per day.

This RAC_{sp} should be compared with the 21-day TWA PEC_{sw}.

2.3. Non-testing methods

Guidance, largely following the European Chemicals Agency (ECHA) recommendations (ECHA, 2008), is provided on the use of non-testing methods in PPP RA, such as (Quantitative) Structure–Activity Relationship ((Q)SAR) models, expert test systems and analogue read-across as tools for deriving intrinsic properties of substances. Non-testing methods may be used to estimate endpoints for metabolites without the toxophore and for impurities. In addition, (Q)SARs might, together with available test data, be used to rank species for identifying the most likely sensitive taxonomic group to focus experimental testing (EFSA PPR Panel, 2012a). For a detailed description of non-testing methods see section 10.1.

Only suitable models (e.g. covering the right domain) with a high predictive reliability should be used (see section 10.1.2). This should, among others, be reflected in the level of statistical significance required for estimates from (Q)SAR models. Validation parameters should ideally indicate good fits (e.g. $Q^2 > 0.7$, concordance correlation coefficient (CCC) > 0.85).[8] Estimates of toxicity should, where possible,[9] be assisted by confidence intervals around the prediction. In case the standard deviation exceeds the predicted value itself, such values should not be accepted. Generally, the worst-case endpoint from several modelling approaches should be used.

Estimates should be confirmed by using weight-of-evidence approaches where all available information is taken into account. This could include a combination of the different (Q)SAR model predictions combined with read-across and other available information like non-standard test data and TK/TD information from mammals.

[8] For further details consult ECHA (2008) guidance.
[9] Not all (Q)SAR models provide standard deviations for predictions.

To date, most experience is gained with (Q)SAR models that predict acute toxicity. It is noted that fewer valid (Q)SAR models are currently available for deriving chronic toxicity data.

A decision scheme for use of non-testing systems is presented below (see section 10.1).

1. Is the (Q)SAR model valid – i.e. is it relevant and reliable (following 5 OECD principles for assessing (Q)SAR models). For example, is the prediction accurate enough (recommended assessment values Q^2, CCC and SD)?

   Yes: Go to 2
   No: (Q)SAR should not be used – consider other model

2. Do the substance and model match – i.e. is the chemical of interest within the scope of the model? In order to do so, the following aspects should be considered:

   - Is the chemical in the applicability domain of the model, as described for the model?
   - Is the substance sufficiently similar to the compounds in the training set of the model (taking structural similarity but also and especially toxophores into account)?
   - Is the prediction for similar substances in the training set reliable (i.e. in line with the experimental data)?

   Yes: Go to 3
   No: (Q)SAR should not be used – consider other model

3. Does model prediction take into account relevant substance properties (e.g. for aquatic toxicity consider water solubility, log K_{ow}, degradability and volatility)?

   Yes: Go to 4
   No: (Q)SAR should not be used – consider other model

4. Are reliable estimations available from more than one (Q)SAR model?

   Yes: Use lowest predicted (Q)SAR endpoint in RA or as qualifier for testing if confirmed by weight of evidence approach

   No: Single value could be used as qualifier for testing if clearly confirmed by weight of evidence approach

2.4. Metabolites and degradation products

The PPR Panel has developed an assessment scheme for RA of metabolites. For the assessment of the metabolite, the applicant has to provide a reasoned case whether the molecule contains a toxophore or if it has been lost following transformation. In case of doubt, it should be assumed that the toxophore remains and that the molecule has a specific mode of action (see assessment scheme below). A first step in this assessment scheme is based on a worst-case approximation of the toxicity of such a metabolite (see section 10.2.7). Testing is required for metabolites with remaining toxophore (see section 10.2.6). The scheme has been developed in order to facilitate the selection of the most appropriate and pragmatic assessment route for metabolites. However, possible endocrine disruption properties should be addressed separately (see section 3.3.6).

1. Is the exposure to the metabolite in the toxicity test with the a.s. measured in the course of the test and adequate for assessing the potential effect of the metabolite (see section10.2.5)?

   Yes: Go to 2
   No: Go to 3

2. Perform the RA assuming all the effects observed in the test with the a.s. can be attributed to the metabolite (see section 10.2.4). Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Go to 3

3. Is it clear that the toxophore has been lost from the molecule (see section 10.2.3 and 10.2.7)?

   Yes: Go to 6
   No or unclear: Go to 4

4. Identify the species or taxonomic group[10] determining the lowest tier 1 $RAC_{sw;ac}$ for the a.s. Is the acute metabolite $L(E)C_{50}$ > 10 times the a.s. $L(E)C_{50}$ (on a molar basis) (see section 10.2.6)?

   Yes: Go to 6
   No: Go to 5

5. Identify the species or taxonomic group[10] determining the lowest tier 1 $RAC_{sw;ch}$ of the a.s. Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Consider higher tier refinement

6. Assume that the acute and chronic[11] toxicity of the metabolite is equal to the toxicity of the a.s. for all first tier taxonomic groups (see section 10.2.7). Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Go to 7

7. Are reliable and adequate non-testing predictions of toxicity (see section 10.2.8 and 10.1) available for all first tier taxonomic groups (fish, plants and invertebrates) for which risks were identified in step 6? Are $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$ using these predictions?

   Yes: Low risk
   No: Go to 8

8. Determine the acute and chronic[10,11] toxicity for those taxonomic groups for which risks were identified in step 6 and/or 7, and where a valid non-testing prediction of toxicity is not available or for which a risk was identified using predicted toxicity. Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Consider higher tier refinement

The principles for assessing metabolites should, in essence, be the same as those for a.s.. However, in contrast to the a.s., data requirements for metabolites do not always have to be addressed by experimental studies (see assessment scheme above and section 10.2.5).

[10] Consider testing with tier 1 sediment organisms if metabolite is distributed in sediment (e.g. *Chironomus* for insecticides and *Lumbriculus* for fungicides).
[11] If chronic RA is triggered by fate properties of the metabolite.

If testing is required for metabolites with the remaining toxophore, it can, in a first step, be limited to the taxonomic group that was identified to result in the lowest tier 1 $RAC_{sw;ac}$ and $RAC_{sw,ch}$ for the a.s. If, however, testing with this taxonomic group shows that this taxonomic group is not sensitive (i.e. the acute endpoint is greater than a factor 10 higher as compared with the parent, on a molar basis[12]) then the RA needs to be continued, assuming that the most sensitive taxonomic group is unknown and the risk to all taxonomic groups should be addressed. If it is unclear whether the toxophore remains and the most sensitive group is not known, then the RA needs to address all taxonomic groups.

If it is clear that the toxophore has been lost from the metabolite, in most cases metabolites are less toxic to the target organisms than the a.s.. As a pragmatic and conservative approach for metabolites without the toxophore the estimates of exposure could be compared with the RAC_{parent} based on the most sensitive endpoint of the a.s. in the relevant compartment. In general, the toxicity needs to be further addressed only if this trigger is failed. For metabolites which have lost the toxophore, the acute and long-term hazard and risk can be addressed using non-testing predictions of toxicity (see further in section 10.1). If the trigger is failed using predicted toxicity then testing is required (see below).

For metabolites (without a toxophore) which require experimental studies (see assessment scheme above), acute toxicity tests with *Daphnia,* rainbow trout and an alga should be conducted. In general, the same testing scheme for a.s. (see Table 22) is required and, hence, only if the metabolite proves to be of similar toxicity to the a.s. should additional species also be tested.

In principle, for metabolites found in the sediment of a water-sediment study, the same triggers for testing should be applied to metabolites as for the a.s. (see section 7.2.5.1). Consequently, when accumulation of a substance in aquatic sediment is indicated or predicted from environmental fate studies, the impact on a sediment-dwelling organism should be assessed. Clearly, the potential to exclude testing on the basis of toxicity will depend on the data available for the metabolite. The applicant should therefore make a case as to whether sediment testing can be waived based on what is known about the fate properties and toxicity profile of the metabolite. For example, if RAs with *Daphnia* indicate that the potential risks are low (taking into account the exposure situation in the sediment), then no further testing should, in general, be required. As a first screening step for metabolites partitioning to the sediment, a formula based on equilibrium partitioning theory, as outlined in the 'TGD part II section 5.5.3' (EC, 2003), can be used to indicate if actual testing is needed. Only if a risk is indicated using this formula should actual testing with sediment organisms be required. This will be further addressed in a PPR Panel opinion on sediment RA under development.

In order to decide whether chronic assessment is necessary, the intended uses and the fate and behaviour of the metabolite should be taken into account. In general, chronic/long-term assessments are required for metabolites where exposure of surface water is likely and the metabolite is deemed to be stable in water, as defined in the data requirements (i.e. there is less than 90 % loss of the original substance over 24 h via hydrolysis under relevant pH conditions (Commission Regulation (EU) No 283/2013)). However, as hydrolysis studies are rarely available for metabolites, the 90 % loss trigger can be applied on data from other abiotic/biotic degradation studies.

For metabolites where chronic testing is necessary, the choice of taxonomic group(s) to be studied for chronic testing should take account of any acute toxicity data on the metabolite. Where information on the acute sensitivity of fish and invertebrates for a particular metabolite is available, chronic testing should only be required on the more sensitive group (i.e. that are a factor of 10 more sensitive). If *Daphnia* is suspected to be insensitive based on the mode of action of the a.s. (e.g. it is an insect

[12] The statement to check whether the LC_{50} of the metabolite is greater than 10 times the LC_{50} of the a.s. on a molar basis means:

$$LC50_{met} > 10\frac{M_{met}}{M_{ai}} LC50_{ai}$$

where $LC50_{met}$ and $LC50_{ai}$ are mass concentrations (mg/L) of metabolite and a.s. at 50 % mortality and M_{met} and M_{ai} are the molar masses (g/mol) of the metabolite and a.s..

growth regulator or a neonicotinoid) then it is necessary to conduct a chronic study using the insect *Chironomus* sp. with the metabolite.

For unstable a.s. (i.e. there is more than 90 % loss of the original substance over 24 h via hydrolysis), it may be more appropriate to conduct chronic studies on the stable metabolite instead of the parent compound. For unstable a.s., where chronic toxicity data for the parent compound are not available and an environmentally significant metabolite exceeds the persistence criteria (i.e. there is less than 90 % loss of the original substance over 24 h via hydrolysis), chronic toxicity data should be submitted for this metabolite regardless of its acute toxicity.

The endocrine disrupting properties of metabolites should also be addressed. However, until common criteria are developed and agreed by the Commission, it is difficult to give specific guidance on how to assess endocrine disrupting compounds (EDC) in relation to PPPs (see section 3.3.6). Therefore, further guidance on how to assess EDCs might be given later when the work of the Commission is finalised. Nevertheless, based on structural properties of the metabolites and also based on information on related compounds indicating that the metabolite may exhibit endocrine disrupting properties, chronic/long-term tests with fish should always be required with this metabolite.

The BCF should be determined as for a.s., if the metabolite is stable (i.e. there is less than 90 % loss of the original substance over 24 h via hydrolysis) and has a log P_{ow}> 3. In the first instance, (Q)SARs could be used to predict the potential BCF. If appropriate information on the bioconcentration/bioaccumulation potential of the metabolite is available from parent BCF data, or other animal metabolism studies, this can be taken into account.

2.5. Combinations of a.s. in formulations (guidance on toxic unit approaches)

The Regulation (EC) No 1107/2009, in Article 29, requires that 'interaction between the a.s., safeners, synergists and co-formulants shall be taken into account' in the evaluation and authorisation. The following guidance is provided to perform the RA for formulations containing more than one active substance. For the details, please refer to section 10.3.

Note, this mixture RA scheme has to be carried out for each endpoint and exposure scenario separately unless it is evident that one specific endpoint/exposure scenario combination clearly drives the risk. The scheme as shown is focusing on effect concentrations (ECx) but may equally be applied to NOEC data if these are pragmatically considered as low effect concentrations.

1. Are measured toxicity data (ECx) available for the given endpoint (typically chronic data available only for a.s.)?

   Only for the a.s. ($ECx_{a.s.}$): Go to 7
   For both formulation (ECx_{PPP}) and a.s. ($ECx_{a.s.}$): Go to 2

2. Check the plausibility of the measured formulation toxicity (ECx_{PPP}) against the calculated mixture toxicity $ECx_{mix\text{-}CA}$ (assuming concentration addition (CA), Equation 13) for exactly the mixture composition of the a.s. in the formulation (ECx_{PPP}) by means of the model deviation ratio (MDR = $ECx_{mix\text{-}CA}/ECx_{PPP}$).

   If MDR = 0.2–5 (CA approximately holds for the mixture): Go to 3
   If MDR > 5 (mixture more toxic than CA): Go to 10
   If MDR < 0.2 (mixture less toxic than CA): Go to 9

3. Check whether the mixture composition in the formulation study giving the measured mixture toxicity (ECx_{PPP}) in terms of the relative proportions of the individual a.s. is similar to the

mixture composition at the PEC_{mix}[13]. As a direct comparison on the basis of the relative proportions of the a.s. at the ECx_{PPP} to the relative proportion at the PEC_{mix} is not informative as such, the comparison is done based on calculated mixture toxicity (assuming CA) for both mixture compositions. Therefore, calculate $ECx_{mix\text{-}CA}$ (see Equation 13) for the mixture composition of the a.s. at the PEC_{mix} and compare with the estimate calculated for the formulation (as already done in step 2 above).

If $ECx_{mix\text{-}CA}$ (a.s. in PPP)/$ECx_{mix\text{-}CA}$ (a.s. in PEC_{mix}) = 0.8–1.2 (mixture similar): Go to 4
If not (mixture not similar): Go to 5

4. Conduct a mixture RA based on measured mixture toxicity, with the exposure–toxicity ratio (ETR_{mix}) being defined as the PEC_{mix} divided by the measured ECx_{PPP} and compare the outcome with the acceptability criterion (trigger value) decisive for the specific endpoint/exposure scenario combination.

   If ETR_{mix} < trigger: Low risk
   If ETR_{mix} > trigger: Low risk not demonstrated/check refinement options

5. Check whether one mixture component clearly drives the toxicity if considering the measured mixture toxicity (ECx_{PPP}), i.e. does the largest part of the sum of toxic units (Equation 14) **calculated for the formulation** (≥ 90 %) come from a single a.s. (toxic unit $(TU)_i$)[14]?

   Yes (single 'driver' of mixture toxicity identified): Go to 6
   No: Go to 8

6. Conduct a RA based on single-substance toxicity data ($ECx_{a.s.}$) for the identified 'driver' of mixture toxicity, with the $ETR_{a.s.}$ being defined as the $PEC_{a.s.}$ divided by the measured $ECx_{a.s.}$ and compare the outcome with the acceptability criterion (trigger value) decisive for the specific endpoint/exposure scenario combination.

   If $ETR_{a.s.}$ < trigger: Low risk
   If $ETR_{a.s.}$ > trigger: Low risk not demonstrated/check single-substance refinement options

7. Is there evidence that synergistic interactions between mixture components might occur (e.g. based on toxicological knowledge from literature or from counter-checking measured and calculated mixture toxicity in other species) which cannot be ruled out for the given species with sufficient certainty?

   Yes (mixture toxicity calculation not feasible): Measured mixture toxicity data required for RA (if becoming available: Go to 2)
   No (mixture toxicity calculation feasible): Go to 8

8. Conduct a mixture RA based on the calculated mixture toxicity according to section 10.3.8:

$$ETR_{mix-CA} = \frac{PEC_{mix}}{ECx_{mix-CA}}$$

[13] Define the mixture to be assessed in terms of the relative proportions (p_i) of the individual mixture components (i) at the PEC_{mix} with p_i being defined as the PEC of the individual components (PEC_i) divided by PEC_{mix}. For an initial screening consider per default the $PEC_{sw\ max}$ of the individual active substances contained in the formulation (i.e. PEC_{mix} equals sum of PEC_i). Additionally check whether metabolites of ecotoxicological relevance have to be included into the PEC_{mix} or not).

[14] With TU_i being defined as the concentration of the *i*th a.s. at the ECx_{PPP} (recalculated to the sum of a.s.) divided by the respective single-substance toxicity ($ECx_{a.s.}$).

If $ETR_{mix\text{-}CA}$ < trigger: Low risk
If $ETR_{mix\text{-}CA}$ > trigger: Low risk not demonstrated, check single-substance refinement options

If the endpoints to be used for the RA refer to the same taxonomic group but are associated with different AFs (e.g. single species test, Geomean or SSD), the calculation of the mixture risk is assessed by:

$$RQ_{mix} = \sum_{i=1}^{n} \frac{PECi}{RACi}$$

If RQ < 1: Low risk
If RQ > 1: Low risk not demonstrated/Check exposure refinement options (see section 10.3.10)

9. Carefully recheck the apparent **antagonism** as observed in the measured mixture toxicity data (ECx_{PPP}) regarding potential impacts of the default assumption of CA and/or heterogeneous input data used for the CA calculation. Does the apparent antagonism remain and is there no toxicologically plausible explanation available (e.g. special feature of the formulation type)?

   Yes (measured mixture toxicity not plausible): Go to 8
   No (measured mixture toxicity plausible): Go to 3

10. Carefully recheck the apparent **synergism** as observed in the measured mixture toxicity data (ECx_{PPP}) regarding potential impacts of heterogeneous input data (a.s.) and of co-formulants ignored in the CA calculation. Does the apparent **synergism** remain?

    Yes: Go to 3. If measured data are not available (see section 7.5.2), or if the assessment in point 3 indicates that the mixtures are not similar, **go to 8** (use modified ETR trigger values, see section 10.3.4)

    No: Go to 3

3. Introduction

3.1. Legislative background

In 2008, the PPR Panel was tasked by EFSA to revise the GDs for Aquatic Ecotoxicology (SANCO/3268/2001 rev 4 final; EC, 2002a) and Terrestrial Ecotoxicology (SANCO/10329/2002 rev. 2 final; EC, 2002b), which were used in the routine RA of PPPs in the context of Directive 91/414/EEC[15]. The replacement of Directive 91/414/EEC by Regulation (EC) No 1107/2009[16] (hereafter the Regulation) in June 2011 called for revision of the existing guidance documents in order to take on board new elements in the environmental RA, e.g. cut-off criteria, biodiversity and new endpoints.

Moreover, the former data requirements (Annex II and III to Directive 91/414/EEC) were revised to be in line with the new Regulation (EC) No 1107/2009. They are now available as Commission Regulation (EU) No 283/2013[17] laying down the data requirements for the dossier to be submitted for the approval of a.s. contained in PPPs and Commission Regulation (EU) No 284/2013[18] for the approval of PPPs. These documents provide information on the basic data requirements for the authorisation of PPPs.

Article 8(5) of Regulation (EC) No 1107/2009 includes a legal obligation to submit scientific peer-reviewed open literature data. EFSA (2011) has provided guidance containing a definition of scientific peer-reviewed open literature. EFSA (2011) also provides guidance on how to identify, select and include scientific peer-reviewed open literature, and how to report the literature search and selection process in a dossier. Regulation (EC) No 1107/2009 and the Uniform Principles (Regulation (EC) No 546/2011[19]) include the decision-making criteria for the approval of a.s., safeners and synergists at the EU level and authorisation of PPPs at the Member State level. A procedure to derive SPGs on the basis of aquatic key drivers and ecological entities to be protected and the magnitude of the tolerable effect (including its spatio-temporal dimension) was proposed by EFSA PPR Panel (2010a) (see also chapter 5).

3.2. Objectives of the Guidance Document

This GD is the first deliverable within the PPR Panel mandate of the revision of the former GD on Aquatic Ecotoxicology (SANCO/3268/2001 rev.4 (final), EC, 2002a), EFSA-M-2009–0001. An overview of the three deliverables under this mandate is given below:

1. Guidance of the PPR Panel on tiered risk assessment for aquatic organisms in edge-of-field surface waters

2. Scientific Opinion of the PPR Panel on the effect assessment for pesticides on sediment organisms in edge-of-field surface waters

3. Scientific Opinion on the state of mechanistic effect modelling approaches for regulatory risk assessment of pesticides for aquatic organisms

[15] Council Directive 91/414/EEC concerning the placing of plant protection products on the market of 15 July 1991, OJ L 230, 19.8.1991, p. 1.
[16] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L 309/1, 24.11.2009, pp. 1–50.
[17] Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, pp. 1–84.
[18] Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, pp. 85–152.
[19] Regulation (EC) No 546/2011 of 10 June 2011 implementing Regulation (EC) No 1107/2009 of the European Parliament and of the Council as regards uniform principles for evaluation and authorisation of plant protection products.

The Guidance presented in this document is intended to replace the current GD on Aquatic Ecotoxicology (EC, 2002a), which was designed to 'provide guidance in the context of the review of a.s. under Council Directive 91/414/EEC […] and not intended to prejudice the authority of MS in national authorisations'. The scientific opinions to be delivered under the mandate will look in detail at some specific aspects that are outlined in the 'Terms of Reference' section above. The further information might be incorporated, if needed, in a future review of this GD.

A revision of the former GD on Aquatic Ecotoxicology (EC, 2002a) became necessary for several reasons, with the main reasons listed below:

- The requirements of the new Regulation (EC) No 1107/2009 and underlying Annexes;
- The new Regulations laying down the data requirements for the dossiers to be submitted for the approval of a.s. contained in PPPs (Commission Regulation (EU) No 283/2013) and for the authorisation of PPPs (Commission Regulation (EU) No 284/2013);
- Recommendations of scientific opinions of the PPR Panel of EFSA with respect to aquatic RA procedures (e.g. EFSA, 2005a, 2006a, 2006b, 2009a; EFSA PPR Panel 2010a);
- Concerns formulated by risk assessors and risk managers of EU Member States and other stakeholders with respect to under- or overprotectiveness of the current aquatic RA procedures and/or RA issues where there is scope for different interpretations (EFSA, 2009b);
- Relevant issues identified during the PPP Risk Assessment Peer Review of a.s.;
- Uncertainties in the current FOCUS surface water scenarios (FOCUS, 2001) and exposure modelling;
- Recommendations of recent workshops (e.g. Ampere, 2007; ELINK, Brock et al., 2010a; AMRAP, Maltby et al., 2010; LEMTOX, Thorbek et al., 2010);
- State-of-the-art new knowledge on the ecotoxicology of PPPs as published in the scientific literature (particularly for PPPs with novel toxic modes-of-action).

This GD on the tiered RA scheme for aquatic organisms in edge-of-field surface waters is intended to provide guidance for applicants and Member States in the context of the authorisation of PPPs and their a.s. under Regulation (EC) No 1107/2009.

3.3. Focus and restrictions of the Guidance Document

3.3.1. Scope of risk assessment

The GD is intended for use in the RA of single substances or combinations of a.s. in one formulation. When a compound or formulation is applied more than once in a growing season, the number of applications is taken into account in the RA.

However, when more than one PPP is applied simultaneously or sequentially within one season, such conditions are not addressed.

3.3.2. Aquatic organisms living in the water column

This GD focuses on a tiered RA procedure for aquatic organisms living in the water column in edge-of-field surface waters. Nevertheless, a tier 1 RA procedure for sediment-dwelling organisms on the basis of the 28-day water-sediment test with *Chironomus riparius* or *Lumbriculus* spp. is also incorporated, since this concerns a data requirement under the PPP regulation. A later PPR scientific opinion in the series mentioned above will deal, in detail, with the effect assessment for sediment-dwelling organisms by paying attention to a wider array of sediment-dwelling species.

3.3.3. Spatial scale: edge-of-field surface waters

Edge-of-field surface water bodies (ditches, streams and ponds) are defined as surface water systems that are as close to a treated field as possible according to agricultural practice. This GD focuses on aquatic organisms in edge-of-field surface waters since appropriate prospective exposure and effect assessment tools for surface waters further downstream that address environmental risks of PPPs at a larger spatial scale (e.g. catchment) are mainly a research activity to date and new tools and knowledge is expected to become available in the next years only. In section 11.5, however, a short description is given on the main differences in RA procedure between the PPP Regulation and the Water Framework Directive (WFD)[20] and on the basic data requirements that should be provided for PPPs by the applicant to derive WFD environmental quality standards.

3.3.4. Use of effect modelling and combination to exposure modelling

Promising effect models at different levels of biological organisation are currently under development in different research projects. Most of these modelling approaches are not yet ready for use in regulatory RA. Further guidance on effect modelling (including TK/TD, population and food web models) and combined effect and exposure modelling approaches will be provided in a future scientific opinion of the PPR Panel of EFSA.

3.3.5. Use of data on marine organisms

The data requirements for PPPs mainly focus on data for freshwater organisms for the effect assessment. Only in some cases are data for marine species (i.e. *Americamysis bahia*) requested. In addition, the exposure assessment is performed for freshwater only, due to the focus on edge-of-field surface waters and the available exposure assessment tools. This GD, however, provides some recommendations on how to use additional toxicity data for marine organisms in combination with toxicity data for freshwater organisms (chapter 8).

3.3.6. Endocrine disruption

As set out in Annex II, Regulation (EC) No 1107/2009 makes specific provision only to approve an a.s., safener or synergist 'if it is not considered to have endocrine disrupting properties that may cause adverse effects on non-target organisms unless the exposure […] is negligible' (section 3.8.2). It follows from this wording that the regulation takes a hazard-based approach to approval where endocrine disrupters are concerned.

Test methods currently required in the data requirement for environmental effects in the PPP Regulation are not designed specifically for identification of endocrine disrupters. However the OECD conceptual framework for testing and assessment of potential endocrine disrupters (OECD, 2012a) includes several test guidelines (in levels 3, 4 and 5) designed for identification of endocrine disrupting properties relevant to wildlife. Of relevance to the current Guidance on aquatic toxicity, these include the fish short-term reproduction assay (OECD TG 229), the 21-day fish screening assay (OECD TG 230), the amphibian metamorphosis assay (OECD TG 231), and the androgenised female stickleback screen (GD 140) at level 3 (*in vivo* tests providing data about selected endocrine mechanisms/pathways), the fish sexual development test at level 4 (*in vivo* assays providing data on adverse effects on endocrine-relevant endpoints), and the fish full life cycle (FFLC) toxicity test (US EPA OPPTS 850.1500) at level 5 (*in vivo* assays providing more comprehensive data on adverse effects on endocrine-relevant endpoints over more extensive parts of the life cycle of the organism). OECD 229 to 231 and 234 are also referred to in the respective Commission Communication[21] (list of

[20] Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. OJ L 327/1, 22.12.2000, pp. 1–72.

[21] Commission Communication in the framework of the implementation of Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market Text with EEA relevance, OJ C 95, 3.4.2013.

test methods and guidance documents relevant to the implementation of Regulation (EU) No 283/2013).

Robust scientific criteria for identification of substances with 'endocrine disrupting properties' are therefore needed to inform updated data requirements in support of this provision of the regulation, and by 14 December 2013 the Commission must present a draft proposal for such criteria. Pending these criteria, substances identified under Regulation (EC) No 1272/2008 (classification and labelling) as C2 and R2, or R2 with evidence of toxicity to endocrine organs, shall be considered to have endocrine disrupting properties, with respect to human health assessment. There is currently no provision for identifying endocrine disruptors within the environmental effects assessment required for the regulation.

In support of future inclusion of such provisions in the PPP Regulation and other pieces of EU legislation, EFSA's Scientific Committee was mandated in 2012 to elaborate a scientific opinion on this issue. This activity occurred in parallel with similar activities initiated by DG Environment (a 'Member States Ad hoc Group on scientific aspects related to identification of endocrine disruptors' (regulatory oriented) and an 'Advisory Group to provide advice to the European Commission on scientific issues relevant to criteria for the identification of endocrine disrupters' (scientific oriented, lead by Joint Research Centre (JRC), Munn and Goumenou (2013)). The scientific opinion, prepared by a Working Group of EFSA's Scientific Committee that ensured involvement of other relevant agencies (e.g. ECHA, European Medicines Agency (EMA)) and the Commissions Scientific Committees SCHER (Scientific Committee on Health and Environmental Risks), SCCS (Scientific Committee on Consumer Safety) and SCENIHR (Scientific Committee on Emerging and Newly Identified Health Risks)), took stock of the following:

- existing scientific criteria for identification of endocrine disrupters;
- what an adverse effect is (explicit in widely accepted definitions of an endocrine disrupter e.g. Weybridge and WHO/IPCS) and what a physiological modulation is;
- whether existing toxicity test methods appropriately cover effects of endocrine a.s.

The conclusions of the opinion (EFSA Scientific Committee, 2013) were that endocrine disrupters should be identified on the basis of evidence of (1) endocrine activity (as might be gathered from read-across, *in silico* methods or mechanistically informative *in vitro* or *in vivo* tests), (2) adverse effect(s) in an intact organism or (sub)populations and (3) a plausible causal relationship between the two. The opinion noted the availability of a number of validated *in vivo* tests for primary endocrine disrupting modalities (oestrogen, androgen, thyroid and steroidogenesis) suitable for the identification of endocrine disrupters in species used for ecotoxicological testing. A respective guidance provided by OECD (2012a) for the interpretation of individual test outcomes with regard to some well-known mechanisms of endocrine disruption might be helpful in this respect. These include several tests in fish and one amphibian test, noted above, which would be relevant to the aquatic ecotoxicity assessment. Nevertheless, the opinion also noted that generic scientific criteria for an 'adverse effect' have not been developed to date, and that decisions on this criteria for endocrine disrupting activity will likely have to be made on a case-by-case basis using expert judgement and weight of evidence. Consequently it is not yet possible to elaborate specific guidance on data requirements and their interpretation/next steps in this GD. The impact of Regulation (EC) No 1107/2009 future provisions for non-approval of endocrine disrupters on the aquatic toxicity effects assessment will therefore be accommodated in future revisions of this GD, and other modules (e.g. terrestrial toxicity), once robust scientific criteria are agreed and available.

3.3.7. FOCUS exposure assessment methodology

Assessment of risks to organisms is always a combination of an effect assessment and an exposure assessment (EFSA PPR Panel, 2010a). This GD is almost exclusively limited to the effect assessment. The current exposure assessment for a.s. approval is based on FOCUS (2001, 2007a, b). The level of protection achieved by the current FOCUS surface water exposure assessment methodology is

unknown since it has not been reviewed during the revision of the GD on Aquatic Ecotoxicology by the PPR Panel of EFSA and exposure assessment goals have not yet been defined for surface water (see also section 5.3 and Appendix A). However, the methodology has been used in regulatory decision making throughout the last years and there is currently no alternative standardised exposure assessment methodology. Therefore, it is assumed that the $FOCUS_{sw}$ methodology will continue to be used until updated or new methods become available and adopted by the Standing Committee on Food Chain and Animal Health (SCFCAH) and will replace the existing tools. $FOCUS_{sw}$ is used for approval of a.s. at the EU level. It is also used in some Member States for product authorisation, but also different exposure assessment procedures may be used.

3.3.8. Chemical and biological monitoring

Although chemical and biological monitoring could be helpful for validating the exposure and effect predictions within the RA framework, this document does not contain guidance for monitoring. Setting up an appropriate monitoring programme requires a clear definition of both the exposure assessment goals and the ecotoxicological protection goals. The exposure assessment goals have, however, not yet been defined by the SCFCAH. Even if the exposure assessment goals would have been defined, validating the current exposure assessment methodology is not possible since it is not known to which percentile of the statistical population of concentrations the exposure predictions correspond (see previous section). It is, however, very difficult to design monitoring programmes such that results can be linked to a specific use of a single a.s. in a specific crop while also excluding other confounding factors. Despite this, the PPR Panel acknowledges the importance of monitoring; for example, for the assessment of persistent substances in soil, a tier 5 monitoring approach was recently proposed by the PPR Panel (EFSA PPR Panel, 2010b). The PPR Panel therefore proposes to develop guidance on monitoring as soon as the current exposure assessment procedure has been reviewed.

3.3.9. Permanent water bodies versus water bodies falling temporarily dry

The aquatic exposure assessment is currently performed using FOCUS surface water tools (FOCUS 2001), which are based on water bodies (ditches, streams and ponds) with a minimum water depth of 30 cm. It can be expected that the annual peak concentration is strongly influenced by this minimum water depth because water depths may be close to the minimum depth during applications of PPPs in spring and summer and the spray drift during these applications may lead to high PECs values for shallow water depths. Therefore, whenever the current exposure assessment methodology will be reviewed, a range of minimum water depths might need to be considered. The ultimate choice of whether such assessment is necessary falls under the responsibility of risk managers.

The standard effect assessment is based on test organisms and studies that are not designed to specifically cover organisms occurring in water bodies falling temporarily dry. No detailed data is available about potential differences in the sensitivity to PPP exposure of those species compared with those used in the standard RA procedure. Temporary small standing water bodies are known to have characteristic plant and animal communities depending strongly on the specific hydrological conditions. A uniqueness is reported for the communities of plants (e.g. with particular protected fern species) and of invertebrates, which often is determined by the absence of fish as predators. Moreover, endangered faunal groups such as amphibians and branchiopod crustaceans might be particularly abundant in these water bodies (European Environment Agency (EEA), 2009). Organisms occurring predominantly in temporary ponds are known to be uniquely adapted to the changing environmental conditions, following strategies like dormancy and dispersal to survive. The sensitivity to toxicants of such organisms, as well as their potential to recover, might be affected by these adaptations (Lahr, 1997).

Based on the data requirements and standard ecotoxicological tests available, this GD predominantly addresses the risk for organisms occurring in permanent edge-of-field water bodies, that is, water bodies that contain water throughout the year.

3.3.10. Active substances with new modes of action

The following guidance and RA schemes are generally recommended for use in the authorisation process of a.s. and formulated PPPs. However, for each compound, it should be carefully evaluated whether the proposed steps are addressing all relevant questions related to the individual properties of the compound under evaluation. Particularly for compounds of new modes of action, a thorough analysis has to be done regarding whether additional or different questions need to be tackled and the scheme proposed here is appropriate. Consequently, this may require that experimental information for all effect tiers is provided the first time a compound with a novel working mechanism is registered.

4. The tiered approach, risk assessment terminology and linking exposure to effects

4.1. Introduction

The aquatic RA procedure for PPPs in edge-of-field surface waters consists of two parts:

- Exposure assessment, where time-dependent concentrations in different compartments of the environment are calculated

- Effect assessment, where the time-dependent environmental concentrations are analysed with respect to possible effects on populations and ecosystems.

Within the authorisation procedure of PPPs in the EU, relevant exposure concentrations in edge-of-field surface waters are obtained by adopting the exposure assessment endpoints described in chapter 5 and by applying exposure scenarios and fate models to derive predicted environmental concentrations (PECs). For prospective exposure assessment, harmonised approaches have been developed that enable the prediction of presumed realistic worst-case exposure concentrations in edge-of-field waters. These are documented in the 'FOCUS Surface Water Scenarios' report (FOCUS, 2001) and are further discussed in chapter 6.

Prospective effects assessment relies on the available SPGs (chapter 5) and relevant ecotoxicological and ecological data. The ecotoxicological data usually concern concentration–response relationships derived from controlled experiments with, for example, standard (chapter 7) and additional aquatic test species (chapter 8) or refined exposure and micro-/mesocosm tests (chapter 9). The ecological data usually relate to the 'target image' of the aquatic community in the relevant surface waters, including ecological traits of the important aquatic species at risk. Examples of ecological scenarios for European streams, ditches and ponds are presented in the ELINK document (Brock et al., 2010a). Assessment factors are usually used to extrapolate the experimental concentration–response relationships in space and time to derive RACs. In principle, this conventional extrapolation approach may be replaced and/or adjusted by appropriate modelling approaches, which will be addressed in a later scientific opinion of the PPR Panel.

4.2. The tiered approach

Ideally, when many scientifically underpinned methods are available and costs are not a limiting factor, environmental RAs can be performed by applying the best-available methods. However, in practice, environmental RAs are not based on an unlimited number of environmental fate and ecotoxicity data but on factors like pragmatism, costs and efficacy. When both pragmatism and science drive the assessment, one can understand the development of tiered systems (Posthuma et al., 2008).

Tiered approaches are the basis of environmental RA schemes that support the registration of PPPs under the PPP Regulation (see, for example, Campbell et al., 1999; EC, 2002a; Boesten et al., 2007; EFSA PPR Panel, 2010a). In this context, a tier is defined as a complete effect or exposure assessment resulting in an appropriate assessment endpoint, for example, PEC or RAC. The concept of tiered approaches is to start with a simple conservative assessment and to only do additional more complex

work if necessary (so it implies a cost-effective procedure both for industry and regulatory agencies). Note, however, that the higher tiers should result in protective RA decisions not in conflict with the SPGs set by the competent authorities. According to Boesten et al. (2007) and Solomon et al. (2008) the general principles of tiered approaches are:

- lower tiers are more conservative than higher tiers;
- higher tiers aim at being more realistic than lower tiers;
- lower tiers usually require less effort than higher tiers;
- in each tier all available relevant scientific information is used;
- all tiers aim to assess the same protection goal.

In short, the tiered system as a whole needs to be (i) appropriately protective, (ii) internally consistent, (iii) cost-effective and (iv) address the problem with a higher accuracy and precision when going from lower to higher tiers (see Figure 2).

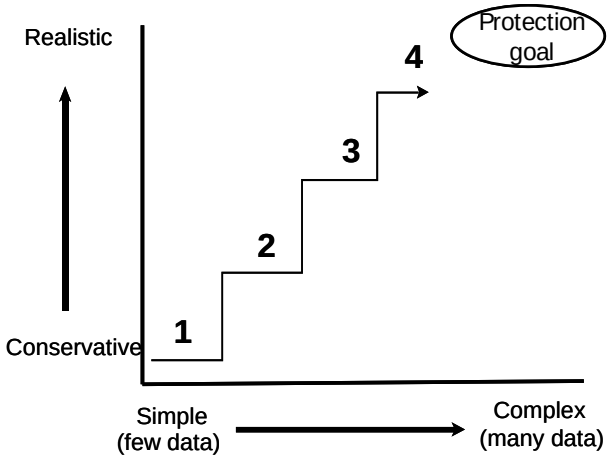


Figure 2: Tiers in the risk assessment process, showing the refinement of the process through the acquisition of additional data (redrafted after Solomon et al., 2008)

An additional practical aspect of the tiered approach is that there has to be some balance between the efforts and the filtering capacity of a tier. For instance, it does not make sense to define a tier that requires 50 % of the efforts of the next higher tier but leads in 95 % of the cases to the conclusion that this next tier is needed (Boesten et al., 2007).

4.2.1. Lower tiers

In PPP RA under the PPP Regulation, the basic data requirements for the first tier RA (set out in Commission Regulation (EU) 283/2013 for approval of a.s.) are strictly defined, in relation to both exposure and effects assessments.

The EU exposure assessment in edge-of-field surface waters normally consists of FOCUS steps 1, 2, 3 or 4 (with the restriction for step 4 that it has to maintain step 3 scenario definitions). These four steps have in common that they can be performed relatively easily and quickly with available tools agreed upon (FOCUS 2001, 2007a; for further details see chapter 6).

For the effects assessment, a logical consequence of the basic dossier requirements is that the RA always starts with the first tier. Detailed information on data requirements for the tier 1 effects assessment in the EU can be found in chapter 7. It should be noted at this point that AFs applied to tier 1 toxicity data to generate tier 1 RACs are specified in the data requirements (see Tables 5.2 and 5.3 in section 7.3).

4.2.2. Higher tier effects assessment

The 'unless'-clauses described in the Uniform Principles (Regulation (EU) No 546/2011) offer the possibility to perform higher tier RAs. Procedures for higher tier effect testing to evaluate the environmental risks of PPPs to aquatic organisms can be found in Chapters 8, 9, 10 and 11 of this GD.

The uncertainties and possible risks indicated by the first tier effects assessment inform the risk assessors and risk managers on which organisms and methods to focus on in the higher tier RA. For example, if the first tier effects assessment for an insecticide indicates that the standard test arthropods are at least an order of magnitude more sensitive than the other standard test species (e.g. algae, fish) the higher tier tests may focus on aquatic invertebrates by performing further studies such as additional laboratory toxicity tests or micro-/mesocosm experiments. It should be noted that if these tests lead to a refined RAC for invertebrates, the risk assessor must then check whether this refined RAC is still protective for other organisms not at risk in the first tier (e.g. fish). Calibration of AFs between higher and lower tier effects assessments is needed.

As noted earlier, with a tiered RA approach, lower tiers should be more conservative than higher tiers. Consequently RACs generated at higher tiers should be higher than those generated by lower tiers – otherwise there is little, if any, incentive to progress from simple and conservative methods to those that more closely resemble the final reference (the actual ecosystem).

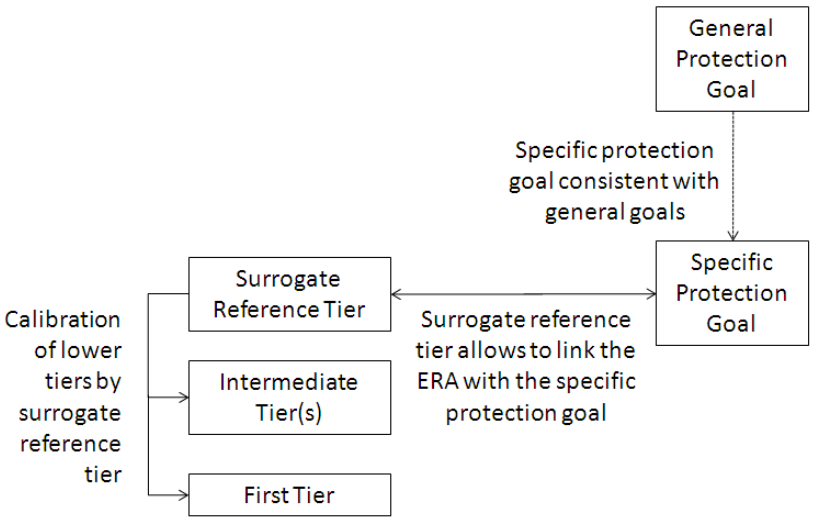


Figure 3: Illustration of the relationship between tiers of the RA process and protection goals in the approach used by the PPR Panel (adapted from EFSA PPR Panel, 2010a). In most cases, assessing directly whether the use of a PPP complies with the SPGs would require refined experimental or modelling methods that would not be practical for routine use in a tier 1 RA procedure. Equally, in general, the standardised studies or models used at a tier 1 level do not measure the SPGs directly. The PPR Panel's solution to this is to identify for each key driver (taxonomic group or other ecological entity) a surrogate reference tier, based on the most sophisticated experimental or modelling RA method currently available that addresses the SPG. This surrogate reference tier will then be used to

calibrate lower tiers using simpler methods that are practical for routine use. Note that from a theoretical point of view the actual ecosystems in the landscape are the final reference.

Another logical consequence of the tiered approach is that higher tiers can be used to calibrate the lower tiers, because the assessment endpoint derived from a higher tier is closer to the actual objectives of the adopted protection goal (Figure 2 and Figure 3). In the aquatic effect assessment for PPPs, an appropriate mesocosm test or model for spatio-temporal extrapolation of toxicity data, in combination with an appropriate AF, is often the highest tier when invertebrates or primary producers are at risk. Appropriate intermediate tiers may be refined exposure studies with standard test species and the SSD approach based on additional toxicity data with potentially sensitive species.

As can be seen from Figure 3 above, it is expected that the surrogate reference tier should link to the SPGs, and meta-analyses of micro-/mesocosm studies provide a scientific underpinning for this assumption (Brock et al., 2006). AFs applied in lower, intermediate tiers (i.e. tier 2A, Geomean and tier 2B, SSD approaches) can then be calibrated against the surrogate reference tier, as described in chapter 8. However, as the AFs for tier 1 are set out in the uniform principles (100 for acute effects on aquatic vertebrates and invertebrates, and 10 for primary producers and chronic effects), the situation may arise where a RAC generated in an intermediate tier using an AF calibrated against the surrogate reference tier is lower than the tier 1 RAC. Evidence from case studies conducted during the preparation of this GD, as well as a comparison of tier 1 RACs with threshold concentrations observed in micro-/mesocosm studies with insecticides presented by Brock and Van Wijngaarden (2012), indicate that this situation occasionally may arise.

This situation indicates that either:

a) the AF applied in the higher tier is too high, that is, there is an error in the calibration of the AF to the surrogate reference tier and/or the linking of the surrogate reference tier to the theoretical real reference (the actual ecosystem); or

b) the AF applied in the first tier is too low.

In the event that a tier 1 RAC is acceptable, there is a possibility that a tier 2 RAC would be lower and the substance should not be approved. In the event that a tier 1 RAC is not acceptable, and higher tier studies (e.g. Geomean, SSD) generate an even lower RAC, it can be assumed that the lower RAC will be used. In the worst case, this could present a disincentive to notifiers conducting higher tier studies, the intention of which is to reduce uncertainty in the RA.

A pragmatic solution to this situation may be to adjust the AF applied to tier 2 data such that the RACs generated in this tier are not lower than the tier 1 RAC. Chapter 8 provides some guidance on the characteristics of tier 2 data that may inform the selection of AFs within the specified range, and the panel recommends that such decisions lie in the realm of risk management. However, this approach evades the more fundamental issue of whether tier 1 AFs, as set out in the uniform principles, are appropriate. The Panel therefore recommends that a future scientific opinion elaborates the scientific justification for the tier 1 AFs.

4.3. Terminology in the aquatic RA of plant protection products

Since the aquatic RA for PPPs follows a tiered approach, characterised by different fate and effect procedures, it is very important to use a transparent RA terminology to facilitate communication between fate experts and ecotoxicologists, between risk assessors and risk managers, and between different stakeholders involved in the administration procedure of PPPs. Before describing the PEC/RAC terminology adopted in this GD, definitions of acute and chronic effect/RAs is provided.

Acute effect assessment = short-term effect assessment

This is the assessment of the RAC for adverse effects of PPP exposure to (non-target) organisms (individuals, populations, communities) occurring within a short period after exposure (hours to weeks; dependent on the life span of the organisms of concern). Note that this is not synonymous with 'assessment of effects due to short-term exposure' since short-term exposure may result in delayed short-term or delayed long-term effects. In current practice of PPP effect assessments, the acute effect assessment scheme starts with the tier 1 acute toxicity dataset. The acute effect assessment may be refined by addressing additional acute toxicity data for the most sensitive taxonomic group triggered by the tier 1 acute core data.

Acute risk assessment = short-term risk assessment

Evaluation/determination of the possibility of adverse effects of PPP exposure to (non-target) organisms in the environment is achieved by comparing the RAC of the acute effect assessment scheme with an appropriate PEC for the environmental compartment of concern. Within the context of the current GD, this PEC will usually be the maximum (peak) concentration derived from the predicted exposure profile.

Chronic effect assessment = long-term effect assessment

This is the assessment of the RAC for adverse effects of PPP exposure to (non-target) organisms (individuals, populations, communities) that develop slowly and/or have a long-lasting course, and that are caused by short-term exposure (latent effects) or long-term exposure. Consequently, a chronic effect assessment is not synonymous with 'assessment of effects due to long-term exposure', but does not exclude it. In current practice of PPP effect assessments, the chronic effect assessment scheme starts with the tier 1 chronic toxicity dataset. The chronic effect assessment may be refined by addressing additional chronic toxicity data for the most sensitive taxonomic group triggered by the tier 1 chronic core data. Since algal tests usually cover the whole life cycle of the test species, the standard 72- to 96-hour toxicity tests with algae can be considered as chronic.

Chronic risk assessment = long-term risk assessment

Evaluation/determination of the possibility of adverse effects of PPP exposure to (non-target) organisms in the environment is achieved by comparing the RAC of the chronic effect assessment with an appropriate PEC for the environmental compartment of concern. Within the context of the current GD, this PEC may be the maximum (peak) or a time-weighted average (TWA) concentration derived from the predicted exposure profile.

Regulatory acceptable concentration (RAC) and predicted environmental concentration (PEC) terminology

In this GD, the PEC derived from a certain exposure assessment approach (e.g. FOCUS scenarios and models) and the RAC[22] derived from different effect assessment tiers will refer to the environmental compartment to which they apply (e.g. surface water (PEC_{sw}; RAC_{sw})).

In addition, the PEC will refer to the type of exposure concentration (peak/maximum or TWA) and the RAC will refer to the type of effect assessment that is addressed (acute (ac) or chronic (ch)). For example:

$PEC_{sw;max}$ the maximum concentration predicted for surface water;

[22] The term RAC was defined by the PPR Panel in EFSA (2006): 'The Annex VI of Directive 91/414/EEC stipulates that an authorisation may be granted if e.g. the predicted short-term exposure does not exceed the concentration of the lowest LC or EC_{50} divided by 100 .i.e., such concentration would be considered acceptable under the regulatory criteria of Annex VI, hence this term'.

$RAC_{sw;ac}$ the regulatory acceptable concentration for surface water within the context of the acute effect assessment scheme.

Furthermore, the effect assessment approach, or the protection goal option addressed, may be added as a preposition when referring to a certain RAC, for example, Geom-$RAC_{sw;ch}$ (the RAC for surface water within the context of chronic effect assessment and derived by means of the geometric mean approach); SSD-$RAC_{sw;ac}$ (the RAC for surface water within the context of acute effect assessment and derived by means of the SSD approach); or ETO-$RAC_{sw,ch}$ (the RAC for surface water within the context of chronic effect assessment and addressing the ETO as derived by means of the model ecosystem approach).

To summarise, the following abbreviations are commonly found in subscript following the term PEC or RAC:

ac: acute
ch: chronic
sw: surface water
max: maximum
twa: time weighted average.

The following prepositions are commonly used before the term RAC:

Tier 1-RAC: RAC on the basis of tier 1 procedure;
Geom-RAC: RAC on the basis of Geomean approach;
SSD-RAC: RAC on the basis of species sensitivity distribution (SSD) approach;
ETO-RAC: RAC on the basis of threshold option and micro-/mesocosms (ETO=ecological threshold option);
ERO-RAC: RAC on the basis of recovery option and micro-/mesocosms (ERO=ecological recovery option).

4.4. Plant protection product effect assessment scheme

From the definitions above, it is clear that, for PPP effect assessment, two distinct effect assessment schemes can be identified that start with the tier 1 acute toxicity data set (acute effect assessment scheme) and the tier 1 chronic toxicity data set (chronic effect assessment scheme).

Since, (1) both the acute and the chronic effect assessment schemes address the same SPG, and (2) the same higher tier effect study (e.g. micro-/mesocosm test or food web model) may be used in both the acute and the chronic RA, the overall effect assessment scheme presented in Figure 4 is a convenient schematic presentation of the tiered approach. A key aspect is that, in PPP effect/RA, both the acute and chronic effects/risks have to be evaluated by starting with the tier 1 approach. The subsequent tiers that follow may differ for the acute and chronic assessment, depending on the remaining uncertainties. On the basis of tier 1 and tier 2 information and appropriate TK/TD models for the species at risk, individual-level effects of time-variable exposure regimes may be assessed. Higher tier data (i.e. tier 2 to tier 4) should never be used in isolation and all available evidence should be taken into account in the RA, ensuring that the higher tier effects data are fully compatible with the tier 1 effect data.

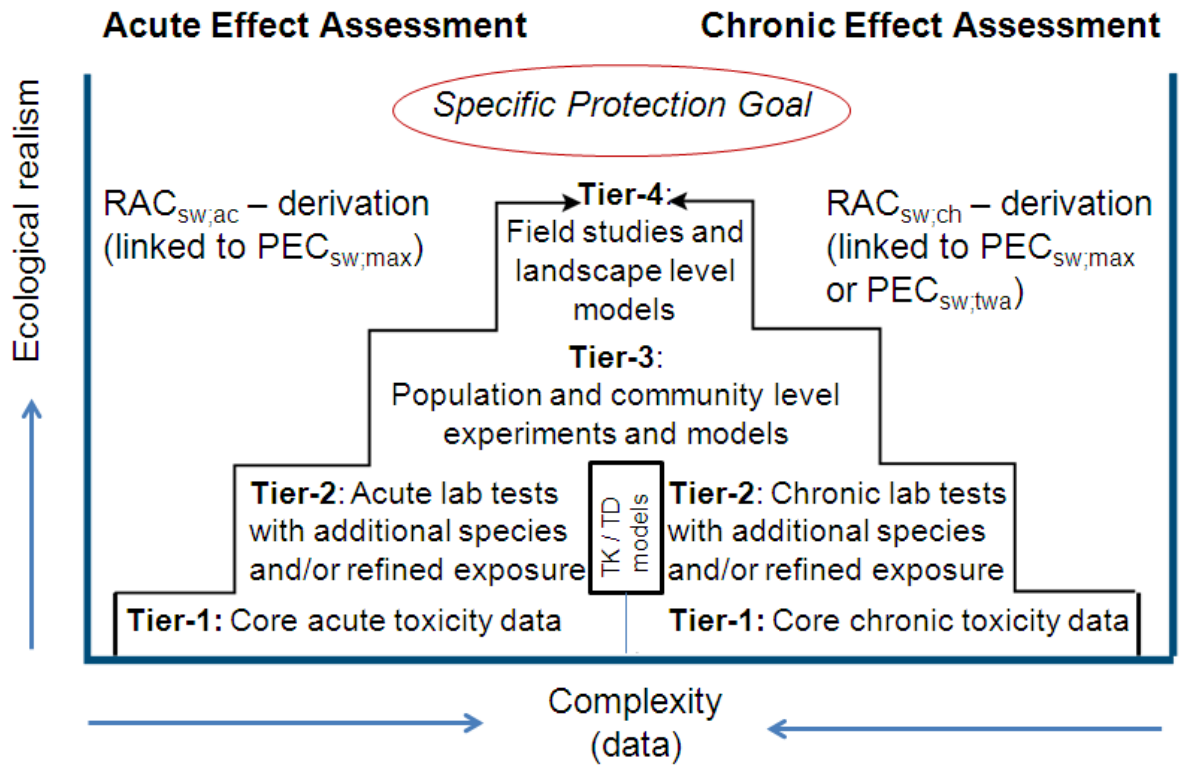


Figure 4: Schematic presentation of the tiered approach within the acute (left part) and chronic (right part) effect assessment for PPPs. For each PPP, both the acute and chronic effects/risks have to be assessed. The tier 1 and tier 2 effects assessments are based on single species laboratory toxicity tests, but to better address risks of time-variable exposures, the tier 2 assessment may be complemented with TK/TD models. Tier 3 (population- and community-level experiments and models) and tier 4 (field studies and landscape level models) may concern a combination of experimental data and modelling to assess population and/or community level responses (e.g. recovery, indirect effects) at relevant spatio-temporal scales. All models included in such a tiered approach need to be properly validated and fulfil required quality criteria

4.5. When to use the peak or a time-weighted average predicted environmental concentration in the risk assessment

A crucial step in the RA is the linking of exposure and effect data. As presented in Figure 4, $PEC_{sw;max}$ values are used in acute RAs, whilst in chronic RAs, in the first instance the $PEC_{sw;max}$ may be used, and under certain conditions a TWA PEC ($PEC_{sw;twa}$) may be used. The use of the TWA concentration approach in the RA of PPPs is based on the observation that effects of PPPs on aquatic organisms may be similar when exposed for a short time to a greater concentration or for a longer time to a smaller concentration, a phenomenon referred to as reciprocity (Giesy and Graney, 1989). Reciprocity relates to Haber's law, which assumes that toxicity depends on the product of concentration and time. For example, an 8-d exposure at 10 µg/L may cause the same effects as a 4-day exposure at 20 µg/L or a 2-day exposure at 40 µg/L, an example of linear reciprocity. Linear reciprocity is the basis of the TWA approach, where exposure concentration is integrated over time (area under the curve (AUC)) and then divided by the duration of the toxicity test. When this approach is applied, different exposure patterns with the same AUC are assumed to have the same effects. Note, however, that for certain PPPs, it has been demonstrated that in prolonged acute toxicity tests, a higher pulse exposure of a shorter duration may be more detrimental than an equivalent lower pulse exposure of a longer duration. For example, Schulz and Liess (2000) demonstrated that after 240 day, a 1-hour pulse exposure to the insecticide fenvalerate was acting substantially stronger on emergence and dry weight biomass of emerged Trichoptera than a 10-hour pulse that was 10 times lower. If short-term, environmentally realistic exposure to PPPs results in delayed responses of PPP-susceptible species (for

other examples see Abel, 1980; Liess, 2002; Brock et al., 2009) this may be a reason to be reluctant in applying the TWA PEC in the chronic RA.

Theoretically, reciprocity should only apply where both uptake and/or elimination of a compound into the test organism (toxicokinetics) and damage and/or repair processes (toxicodynamics) have reached a steady state (Rozman and Doull, 2000). Ashauer et al. (2007) found that for the insecticide chlorpyrifos and the crustacean *Gammarus pulex*, the TWA approach based on an acute toxicity test greatly underestimated mortality in longer term exposure studies, whereas it overestimated mortality caused by pentachlorophenol. In long-term toxicity tests that addressed survival/mortality of *Gammarus pulex*, however, they demonstrated that the TWA concentration approach can be used to extrapolate results of a chronic pulse test to other chronic exposures for both chlorpyrifos and pentachlorophenol. This observation supports the use of the TWA concentration approach in chronic RAs. Moreover, for the herbicide metsulfuron-methyl and the aquatic vascular plant *Myriophyllum spicatum*, it was demonstrated that the TWA approach can be used to assess concentration–response relationships under longer term time-variable exposure conditions (Belgers et al., 2011). In general, the longer duration of chronic tests implies a greater probability that toxicokinetics and toxicodynamics will approach a steady state by the end of the study period.

Although current scientific knowledge is insufficient to generally support the use of the TWA approach in acute RAs (i.e. using a PEC_{twa}), this does not mean that the concentrations of the acute toxicity estimate (the C in the LC_{50} or EC_{50}) should not be expressed in terms of measured TWA. In general, OECD test guidelines recommend that toxicity endpoints (e.g. EC_{50}) should be based on mean measured concentrations. However, if evidence is available to demonstrate that the measured concentration of the test substance has been satisfactorily maintained within ± 20 % of the nominal throughout the test, then the results can be based on nominal.

4.5.1. When and how (not) to use the $PEC_{sw;twa}$ in chronic risk assessments

In chronic RAs, the default recommendation is to use the $PEC_{sw;max}$, and under certain conditions, a TWA PEC may be used. However, $PEC_{sw;twa}$ should not be used if the following conditions apply (adapted after Brock et al., 2010a):

- In chronic RAs that use RACs from effect studies where the exposure has not been maintained, loss of the a.s. in the test system was relatively fast and the toxicity estimate has been expressed in terms of nominal or initially measured concentration. This may, for example, be the case in the 72-hour standard test with algae and the 28-day water-spiked test with *Chironomus riparius*;

- When the effect endpoint in the chronic test (used to derive the RAC) is based on a developmental process during a specific sensitive life cycle stage that may last a short time only (e.g. malformations during metamorphosis, effects caused by endocrine disruption) and evidence exists that the exposure may occur when the sensitive stage is present;

- When the effect endpoint in the chronic test (used to derive the RAC) is based on mortality occurring early in the test (e.g. in the first 96-hours), or if the acute to chronic ratio (acute EC_{50} or LC_{50}/chronic NOEC) both based on immobility or mortality is < 10. If the acute to chronic ratio is small (< 10) the time to onset of maximum effects is likely to be short in the chronic test;

- If latency of effects (delayed effects) has been demonstrated in longer term toxicity tests in which observations continue after the exposure is completed or the organisms have been removed from the stressor, or when latency might be expected on the basis of other data such as toxic mode of action. In longer term studies, latency may result from delays in the chain of events between exposure and expression of effects (e.g. in the case of moulting inhibiting insecticides and substances suspected of endocrine disrupting effects). To demonstrate latency, it may even be required to make observations on the responses of the offspring. It is advised to address latency if, through analogy to similar substances or knowledge of mechanisms of action, it is expected to

occur. In cases where latency is known not to occur in PPPs with a similar toxic mode of action, it might be disregarded. Further information on refined exposure laboratory toxicity tests as a means to address latency is given in section 9.2.

In cases other than those listed above, the use of the TWA approach in chronic RA may be an option. Ecotoxicologists must determine, based on knowledge of ecotoxicological data, whether or not the TWA concentration approach is appropriate to be used in the chronic RA, and which time window the TWA should be based upon.

For realistic to realistic worst-case RA approaches, the time window of the TWA PEC should be equal to or smaller than the length of the relevant chronic toxicity test (or life stage of the species with the highest ecotoxicological concern) that triggered the risk. For invertebrates, fish and macrophytes, a default 7-day TWA time window is proposed if the TWA concentration approach is deemed appropriate (see criteria above) and no further information on the relation between exposure pattern and time-to-onset of the relevant effect is provided. This default value was proposed at the ELINK workshop and is based on expert judgement and pragmatism. For the time being, the PPR Panel of EFSA adopts this pragmatic approach that most likely is relatively worst case. It may be justified to lengthen or shorten the default 7-day TWA period when scientific data are made available that demonstrate that another TWA period is more appropriate. This, for example, may be demonstrated by means of tailor-made experiments that allow the comparison of the effects of different exposure durations (including the onset-of-effects) on the organisms of concern and/or TK/TD modelling (e.g. Jager et al., 2011).

If the use of the TWA approach in the chronic RA is appropriate, concentration–response relationships observed in toxicity tests with long-term exposure (which may be variable in time), as well as the derived RAC, can be expressed in terms of TWA concentrations. This RAC value can be compared with the appropriate TWA PEC under the condition that the time window for the PEC estimate is equal to or shorter than that for the 'C' in the effect estimate.

Further information on the use of TWA in combination with endpoints derived from experimental ecosystems is given in section 9.3.5.2.

4.5.2. Decision scheme to use the $PEC_{sw;max}$ or $PEC_{sw;twa}$ in the risk assessment

For an appropriate RA of PPPs in a relevant edge-of-field surface water, the minimum set of exposure estimates required are the $PEC_{sw;max}$, the highest $PEC_{sw;7d\text{-}twa}$ and the annual exposure profile from which exposure characteristics like number of pulse exposures, pulse durations, intervals between pulses and water dissipation DT_{50} values can be deduced. Furthermore the $PEC_{sw;21d\text{-}twa}$ is a basic data requirement to assess the risks for secondary poisoning (see section 7.6.3). The following decision scheme may help to determine whether the $PEC_{sw;max}$ or the $PEC_{sw;twa}$ can be used in the chronic RA.

1. *Chronic Assessment*. Is $PEC_{sw;max}$ (of highest available tier) > $RAC_{sw;ch}$ (of highest available tier)?

   Yes: Go to 2
   No: Low chronic risk

2. Is the $RAC_{sw;ch}$ derived from a test with algae, or from a long-term (≥ 7 days) test with another water organism and the following conditions apply: (i) loss of the a.s. from water is more than 20 % of nominal at the end of the exposure period and (ii) the toxicity estimate (e.g. EC_{10} or NOEC) is expressed in terms of nominal/initially measured concentration of the a.s.?

   Yes: $PEC_{sw;twa}$ not appropriate (low risk not demonstrated)
   No: Go to 3

3. Is the $RAC_{sw;ch}$ based on treatment-related responses of the relevant test species early in the chronic test (e.g. during the initial 96-hours observed mortality/immobility in tests with animals, or 50 % reduction in growth rate in tests with macrophytes, in the treatment level above the one from which the $RAC_{sw;ch}$ is derived) or is the acute to chronic ratio (acute $L(E)C_{50}$/chronic NOEC or acute $L(E)C_{50}$/chronic EC_{10}) based on immobility or mortality < 10?

   Yes: $PEC_{sw;twa}$ not appropriate (low risk not demonstrated)
   No: Go to 4

4. Is it demonstrated by the notifier that, for the organisms and the PPP under evaluation and/or PPP with a similar toxic mode of action (read-across information), the following phenomena are not likely: (i) latency of effects due to short-term exposure; (ii) the co-occurrence of exposure and specific sensitive life stages that last a short time only?

   Yes: Go to 5
   No: $PEC_{sw;twa}$ not appropriate (low risk not demonstrated)

5. Is $PEC_{sw;7d\text{-}twa}$ (of highest available tier) > $RAC_{sw;ch}$ (of highest available tier)?

   Yes: Go to 6
   No: Low risk demonstrated

6. Are experimental (or TK/TD modelling when guidance is available) data available that demonstrate that for the species a larger time window for the $PEC_{sw;twa}$ may be used (not exceeding the duration of the tier 1 chronic test that triggered the risk)?

   Yes: Go to 4 and replace the $PEC_{sw;7d\text{-}twa}$ by another appropriate $PEC_{sw;twa}$
   No: Low risk not demonstrated

5. Exposure assessment goals and specific protection goals for water organisms

5.1. Introduction

A RA scheme that addresses an SPG requires a clearly defined ERC that needs to be consistently used in both the exposure and the effect assessment procedures within the same RA scheme. Before defining in greater detail the SPGs for water organisms in edge-of-field surface waters, the relevant ERCs and exposure assessment goals will be described and discussed.

5.2. The ecotoxicologically relevant concentration (ERC)

Lack of a clear conceptual basis for the interface between the exposure and effect assessment may lead to a low overall scientific quality of the RA. This interface is defined by EFSA (2005a) and Boesten et al. (2007) as the concentration that gives an appropriate correlation to ecotoxicological effects, and is called the ecotoxicologically relevant type of concentration (ERC). In the RA, the ERC needs to be consistently applied so that field exposure estimates (PECs) and RACs can be compared as readily as possible. The ecotoxicological considerations determining the ERC may include the following questions:

- In which environmental compartment do the aquatic organisms at risk live (e.g. water or sediment)?
- What is bioavailable for the organism (e.g. for pelagic organisms the freely dissolved fraction in water, the sorbed fraction to dissolved organic matter or suspended particulate matter, or the fraction in the food)?
- What is the influence of the time-variable exposure pattern on the effects (e.g. do peak or longer-term concentrations explain the responses)?
- If latency of effects is of concern or longer term exposures explain the responses, what should be the relevant time window of the exposure and effect estimates?

In ecosystems, the ERC may be different for substances that differ in toxic mode of action and for different populations of aquatic organisms, life stages of species, and so on. For example, for an aquatic invertebrate living associated with macrophytes in shallow freshwater ecosystems, the ERC could be the maximum concentration over time of the dissolved fraction for a fast-acting insecticide or some TWA concentration for a slow acting fungicide (e.g. 7-day or 21-day TWA). For detritivores that predominantly dwell on the sediment surface and process particulate organic matter (POM), the ERC for a hydrophobic substance could be the concentration of the PPP in the POM consumed. For an aquatic insect that predominantly dwells at the water surface (e.g. water striders) the ERC of a fast acting insecticide may be the water concentration in the top layer of the water column, which may be relevant if stratification of the insecticide occurs initially.

After the ERCs for the PPP under evaluation and the aquatic organisms at risk have been determined, the collected exposure data can be linked to the relevant ecotoxicological data. It is important that, within the same tiered RA scheme (addressing the same SPG), the type of ERC used to express the 'C' in the PEC estimates should not be in conflict with the ERC used to express the 'C' in the RAC estimates, in the sense that a realistic to worst-case RA can be performed.

This GD focuses on water organisms that dwell in the water column of edge-of-field surface waters. Within the context of this GD the concentration of the freely dissolved chemical (hence, not including

chemical sorbed on suspended matter or sediments[23]) averaged over the water column of edge-of-field surface waters is chosen as the most relevant ERC for toxicity of a PPP in water.

The PPR Panel is aware of the fact that the emission route run-off may lead to a high concentration of the a.s. in suspended particles, a fraction that may be bioavailable to filter feeders. Currently, no appropriate standardised lower tier test is mentioned as dossier requirement to address this issue. In principle, however, the potential effects of particle-associated PPP emission to surface water can be calculated via the FOCUS surface water tools and can be addressed in higher tier effect studies (e.g. mesocosms). Mechanistic effect models developed for pelagic filter feeders may be of help in the near future to better integrate exposure estimates to experimental effect data. It is anticipated that future guidance on the environmental risks of PPPs to sediment organisms will address the relative contribution of aqueous and food exposure in the ERA of PPPs, that also may be of use to develop ERA procedures for pelagic filter feeders.

5.3. Exposure assessment goals in edge-of-field surface waters

The specification of exposure assessment goals has a significant influence on the overall level of protection to be achieved by the proposed RA procedure. However, as described in section 3.3.7, the PPR Panel assumes that the current FOCUS exposure assessment for edge-of-field surface water systems will continue to be used until updated or new methods become available and adopted by the SCFCAH. Therefore, no proposals for the specification of the exposure assessment goals are provided for the time being, instead, a description of the related choices in the current $FOCUS_{sw}$ methodology is provided in Appendix A.

5.4. Specific protection goals for water organisms

The Regulation requires a high level of protection (e.g. no unacceptable effects on the environment (preambles 8, 10, 24 and Article 4.3)). However, only general protection goals are given in the legislation and SPGs that are needed for RA are not precisely defined. In EFSA PPR Panel (2010a), a process is described for defining SPG options for key drivers (main groups of organisms) covering ecosystem services which could potentially be affected by PPPs. The aquatic key drivers, their ecological entities to be protected and related tier 1 test species are mentioned in Table 12. In general, to ensure ecosystem services, taxa representative for aquatic key drivers identified need to be protected at the population level. However, it is proposed to protect aquatic vertebrates (fish, amphibians) at the individual (in the acute RA to avoid visible mortality) to population level (chronic RA). To protect the provisioning and supporting services provided by microbes, it is proposed to protect them at least on the functional group level (EFSA PPR Panel, 2010a).

Table 12: The aquatic key drivers and their ecological entity to be protected as proposed in EFSA PPR Panel (2010a) and the current standard aquatic test species related to these key drivers (Commission Regulation (EU) 283/2013)

Key driver	Ecological entity to be protected	Tier 1 taxa mentioned in data requirements (Commission Regulation (EU) 283/2013)
Aquatic algae	Populations	Green algae, e.g. *Pseudokirchneriella subcapitata* Other taxonomic groups, e.g. the diatom *Navicula pelliculosa*
Aquatic vascular plants	Populations	Monocots, e.g. *Lemna gibba/minor, Glyceria maxima* Dicots, e.g. *Myriophyllum*
Aquatic invertebrates	Populations	Crustaceans: *Daphnia magna/pulex*, *Americamysis bahia* Insects: *Chironomus riparius* Oligochaets*: Lumbriculus* spp.

[23] If the risk is predicted to be via the sediment in a higher tier RA then this should not be ignored; however, detailed guidance will not be provided in this GD.

Key driver	Ecological entity to be protected	Tier 1 taxa mentioned in data requirements (Commission Regulation (EU) 283/2013)
Aquatic vertebrates	Individuals (in acute risk assessment to avoid visible mortality) – populations (in chronic risk assessment)	Fish, e.g. *Oncorhynchus mykiss*
Aquatic microbes	Functional groups	No standard test species

The ultimate goal of the update of the AGD is to develop and describe protective RA schemes based on SPG options for aquatic key drivers (main taxonomic groups to be protected) defined in terms of various dimensions: degree of certainty (which always needs to be high), ecological entity (see Table 13), attribute, magnitude, temporal scale and spatial scale. As mentioned in EFSA PPR Panel (2010a), interdependency exists between the different SPG 'dimensions'. The dimension 'attribute' is closely linked to the dimension 'ecological entity'. For example, functional groups (ecological entity) are often linked to processes (attribute), populations (ecological entity) to abundance/biomass (attribute) and individuals (ecological entity) to behaviour/survival/growth (attribute). Similarly, in edge-of-field surface water the dimension 'magnitude' is closely linked to the dimension 'temporal scale'; for example, a larger magnitude of effect may be acceptable only if the response is short-term and not propagating to the community or a small magnitude of effects may be considered unacceptable if it is long-term.

5.5. Specific protection goal options for aquatic key drivers in edge-of-field surface water

For key drivers in edge-of-field surface waters that need to be protected at the population level or higher, this GD will present assessment schemes that allow derivation of RACs on the basis of two options:

1. Accepting only negligible[24] population-level effects (ecological threshold option, ETO). The reasoning for this approach is based on the consideration that by not accepting population-level effects on representative sensitive populations in edge-of-field surface waters, these populations will be protected and propagation of effects to the community-, ecosystem- and landscape-level will not occur.

2. Accepting some population effects if ecological recovery takes place within an acceptable time-period (ecological recovery option, ERO). When performing a protective RA it is nevertheless desirable to not be overly protective. However, when including recovery to identify (un)acceptable effects, all relevant processes that determine population viability and the propagation of effects to the community-, ecosystem- and landscape-level are to be considered. Only such an integrative assessment can ensure a protective RA. For example, if a temporal reduction of an invertebrate species of some months is accepted, it has to be ensured that organisms preying on this invertebrate can use other adequate food sources that are sufficient to sustain the population of the predator. In addition, if recovery of populations of short-cyclic water organisms is predicted, it also has to be ensured that species with contrasting life cycle traits (i.e. longer generation time) are able to completely recover in the time available between the exposure events. Furthermore, the Regulation requires that the RA methodology should account for the simultaneous use of PPPs (applied in tank mixtures or used in sequence) and that the use of PPPs does not have any long-term repercussions for the abundance and diversity of non-target species (see EFSA PPR Panel, 2010a). The selection of option 1 (ETO), above, for the RA of individual PPPs is more likely to avoid stress caused by the multiple use of different PPPs. Although, a RA that considers recovery of sensitive populations may be a reasonable option for surface waters adjacent to crops with a limited PPP input, it is more uncertain if option 2, SPG (ERO) can be achieved when assessing risks for individual PPPs for their use in crop protection programmes characterised by intensive

[24] The term negligible is used since it is difficult to demonstrate that no effect is occurring.

PPP use (simultaneous or repeated use of different PPPs). However, to draw a meaningful conclusion requires a thorough analysis of PPP use in major crops and the identification of those crops where the ecological risks are unacceptably high due to multi-stress by different PPPs (including the recovery of potentially vulnerable populations). No definitive answer can be given at present and it is recommended to derive and report, whenever possible, both RAC options. Guidance to address combinations of a.s. in formulations is provided in section 10.3 of the AGD.

Below, SPG options are provided for the different aquatic key drivers based on the procedure described in EFSA PPR Panel (2010a) and the two options described above (Table 13 and Table 14 and following text). Note that in this GD the dimension 'spatial scale' is fixed (edge-of-field surface water) and the dimension 'degree of certainty' always should be high. Consequently, the SPGs for aquatic key drivers in edge-of-field surface waters are defined in terms of the dimensions 'ecological entity', 'attribute', 'magnitude' and 'temporal scale'.

Table 13: Overview of proposed specific protection goals for the ecological threshold option

Organism group	Ecological entity	Attribute	Magnitude	Time
Algae	Population	Abundance/biomass	Negligible effect	Not applicable
Aquatic plants	Population	Survival/growth Abundance/biomass		
Aquatic invertebrates	Population	Abundance/biomass		
Vertebrates	Individual	Survival		
	Population	Abundance/biomass		
Aquatic microbes	Functional group	Processes (e.g. litter break down)	RA will not be developed since tier 1 data requirements are not defined	

Table 14: Overview of proposed specific protection goals for the ecological recovery option

Organism group	Ecological entity	Attribute	Effect allowable on most sensitive/vulnerable population	
			Magnitude	Duration
Algae	Population	Abundance/biomass	Small effect[(a)] Medium effect[(a)] Large effect[(a)]	Months Weeks Days
Aquatic plants[(b)]	Population	Survival/growth Abundance/biomass	Small effect[(a)] Medium effect[(a)]	Months Weeks
Aquatic invertebrates[(b)]	Population	Abundance/biomass	Small effect[(a)] Medium effect[(a)] Large effect[(a)]	Months Weeks Days
Vertebrates	No recovery option			
Aquatic microbes	Functional group	Processes	RA will not be developed since Tier 1 data Requirements are not defined	

(a): None of the direct effects should lead to unacceptable indirect effects.
(b): The recovery option will often not be applicable in case organisms with a long life cycle could be affected and short-term (days) large effects generally will be acceptable only for short-cyclic organisms that have a high reproduction capacity. Consequently, strict criteria for (not) allowing the recovery option are given in the further guidance below.

5.5.1. Specific protection goal proposal for algae (e.g. green algae, diatoms, blue-greens) in edge-of-field surface water

For both SPG options, algae will be protected at the population level by considering their abundance/biomass in edge-of-field surface waters. The population level is proposed instead of the functional group level as indicated in EFSA PPR Panel (2010a) since clear definitions of functional groups of algae are lacking. Option 1 (ETO) allows negligible effects on these endpoints only.

Option 2 (ERO) allows large effects for days, medium effects for weeks, and small effects for several months on the abundance and/or biomass of vulnerable populations of algae as long as these effects do not cause persistent indirect effects on other water organisms that depend on algae. In option 2, the range in acceptable magnitude of effects from small to large is selected because large temporal changes in abundance/biomass of algal populations are common in (non-stressed) aquatic ecosystems, due to the short response time to fluctuating environmental conditions such as light and nutrient availability and predation by zooplankton. The required assessment endpoints for algae can be adequately studied in micro-/mesocosms characterised by conditions close to natural in terms of competition, predation and natural stressors. When these conditions are met, such test systems may be considered as a surrogate reference tier to calibrate the lower effect assessment tiers (for details see chapter 9).

5.5.2. Specific protection goal proposal for aquatic vascular plants (e.g. dicotyledonous, monocotyledonous) in edge-of-field surface water

Aquatic vascular plants will be protected at the population level by considering their growth and/or abundance/biomass in edge-of-field surface waters. Option 1 (ETO) allows negligible effects only. Option 2 (ERO) allows medium effects as long as the duration of the effect on the abundance and/or biomass of vulnerable populations of macrophytes is not longer than weeks or small effects when they last for a few months. In option 2, the acceptable magnitude of effects is small to medium since large effects are not desirable even if recovery can be demonstrated. Macrophytes play important ecological roles (e.g. as substrate, shelter, food source) on which many other water organisms depend. As suitable surrogate reference tier for aquatic vascular plants, micro-/mesocosm test systems can be used. According to Maltby et al. (2010) the required assessment endpoints for vascular plants need to be studied in conditions close to natural in terms of competition, predation and natural stressors in order to obtain realistic assessment endpoints. In micro-/mesocosms, however, usually a few free-growing (volunteer) species are dominating the macrophyte community, allowing the study of concentration–response relationships for a few macrophyte populations only. Introducing potted plants in micro-/mesocosms allows the study of the responses of a larger number of macrophyte species. The disadvantage of potted plants is that below-ground competition between species is excluded. It may be an option to combine the two approaches by reserving a part of the mesocosm test system for potted plants and another part for volunteer species. Mechanistic models (e.g. TK/TD models) and/or refined exposure tests with selected species may be used as complimentary tools to address effects of realistic time-variable exposures.

5.5.3. Specific protection goal proposal for aquatic invertebrates in edge-of-field surface water (e.g. crustaceans, rotifers, insects, oligochaete worms, molluscs)

Aquatic invertebrates will be protected at the population level by considering their abundance and/or biomass in edge-of-field surface waters. Option 1 (ETO) allows negligible effects on these endpoints only. Option 2 (ERO) allows small effects for a few months, medium effects for weeks and large effects for days on the abundance and/or biomass of vulnerable populations of invertebrates, as long as their reduction does not result in more persistent indirect effects. In option 2, the range in acceptable magnitude of effects from small to large is selected because large temporal changes in abundance/biomass of particularly short-cyclic invertebrate populations (e.g. daphnids, rotifers, and representatives of oligochaete worms and insects) are common even in more or less pristine aquatic ecosystems. The required assessment endpoints for aquatic invertebrates can be adequately studied in micro-/mesocosms (surrogate reference tier) if the conditions in these test systems are sufficiently representative of natural ecosystems in terms of species composition, species interactions (competition, predation) and natural stressors, and the duration of the experiment is long enough to enable detection of delayed effects. It should be carefully evaluated whether representatives of potentially sensitive and vulnerable species (e.g. uni-/semivoltine insects with a long life cycle) are present. To extrapolate results of micro-/mesocosm experiments, it may be an option to also use population models (not described in this GD) to better address recovery potential of vulnerable invertebrates. These models, however, should consider how recovery is affected by possible interference with other populations (species interactions such as predation and competition).

5.5.4. Specific protection goal proposal for aquatic vertebrates in edge-of-field surface water (e.g. fish, amphibians)

Since mortality of fish and amphibians due to acute toxicity of PPPs is generally not accepted by risk managers, and chronic effects on their populations should not be larger than negligible, the ETO only is proposed as the SPG for aquatic vertebrates. A well-established and widely accepted surrogate reference tier to calibrate the acute and chronic tier 1 effect assessment for vertebrates is not (yet) available. A possible surrogate reference tier for the ETO seems to be the SSD approach (based on acute toxicity data to assess acute effects and on chronic toxicity data to assess chronic effects).

5.6. Vulnerable species

An important follow-up step in the EFSA approach to define SPGs (EFSA PPR Panel, 2010a) is the identification of vulnerable representatives for each aquatic key driver mentioned in section 5.5. In the aquatic RA for PPPs, the identification of vulnerable species is important for (1) designing and interpreting higher tier experiments (e.g. micro-/mesocosm tests), (2) identifying focal species that need special attention when constructing ecological scenarios and adopting mechanistic modelling approaches, and (3) designing and interpreting biomonitoring programmes to evaluate the appropriateness of the prospective RA procedures.

Vulnerability has been defined as 'the degree to which a system, subsystem, or system component is likely to experience harm due to exposure to a hazard' (Turner et al., 2003). Properties relevant to define vulnerability are species traits and characteristics that determine (1) susceptibility to exposure, (2) toxicological sensitivity, and (3) internal and external (recolonisation) recovery processes (see Figure 5).

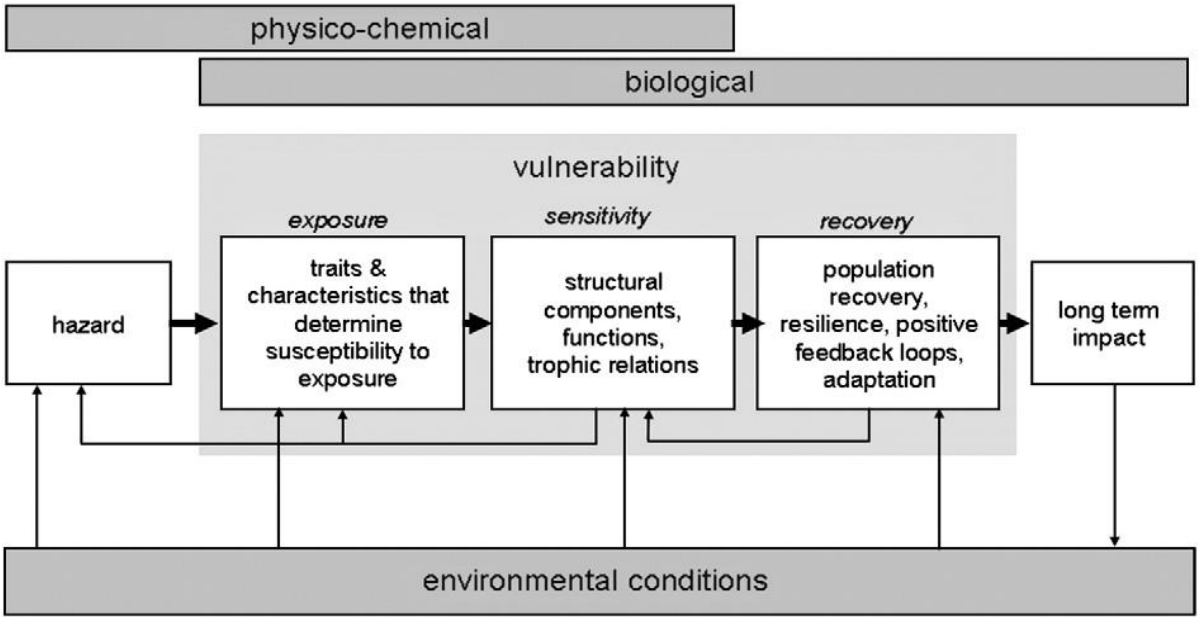


Figure 5: General framework for ecological vulnerability assessment for hazard. Taken from De Lange et al. (2010).

5.6.1. Susceptibility to exposure

Species traits that have been shown to increase exposure include habitat preference in water-sediment or air-water boundary layers of aquatic ecosystems where higher concentrations of PPPs may occur (e.g. Brock et al., 2010b), or selective consumption of food with a high PPP content (e.g. carnivores high in the food chain; Stäb et al., 1996). Examples of ecological traits that reduce exposure are emergence of insects before PPP exposure (Liess et al., 2008) and ability to actively migrate to non-exposed patches of habitat or ecosystems (e.g. Lahr, 1997; Liess and Von der Ohe, 2005).

5.6.2. Toxicological sensitivity

Inter-specific variation in toxicological sensitivity of aquatic species to PPPs has been widely documented (e.g. Von der Ohe and Liess, 2004; Luttik et al., 2011) and partly can be attributed to the specific toxic mode of action of the PPP of concern. For example, in the case of insecticides, aquatic arthropods usually are most sensitive (Maltby et al., 2005) while algae and/or macrophytes tend to be sensitive in the case of herbicides (Van den Brink et al., 2006). Many fungicides, however, have general biocidal properties in that representatives of different taxonomic groups may be categorised as potentially sensitive (Maltby et al., 2009). Recently efforts are being made to mechanistically link biological traits with intrinsic sensitivity of individual species of concern (Rubach et al., 2012). Experimental results indicate that this approach has promise, but, according to Baird and Van den Brink (2007) effort is needed to compile the trait information of species to increase the power, precision and taxonomic representativeness of this approach. Note that toxicological sensitivity of water organisms to a specific PPP may be influenced by the co-occurrence of other stress-factors (e.g. chemical stressors such as other PPPs, physicochemical stressors like low oxygen levels and biotic stressors like pathogens) (Heugens et al., 2001).

5.6.3. Ecological recovery

As regards the recovery of aquatic organisms from PPP stress, a distinction can be made between internal and external recovery processes (Caquet et al., 2007; Brock et al., 2010b). Internal recovery depends on surviving individuals in the stressed ecosystem (e.g. in refuges), or on a reservoir of resting propagules (e.g. eggs) that remain unaffected by the PPP. In particular, species with a short generation time and a high reproductive output are able to recover quickly from PPP stress if a small fraction of the population survives. Laboratory investigations under optimum conditions indicate that internal recovery needs at least one generation time (Barnthouse, 2004). However, field investigations indicated that recovery requires around three generation times under realistic conditions (Niemi et al., 1990; Liess and Von der Ohe, 2005; Kattwinkel et al., 2012). External recovery depends on the immigration of individuals from neighbouring ecosystems by active (e.g. flying) or passive (e.g. wind) dispersal. Species with a strong capacity to actively migrate from one site to another are especially good at performing this process within a relatively short period of time. However, ecological recovery of sensitive species may be hampered by other less sensitive and competitive species that increased subsequent to PPP application, or by physicochemical environmental conditions that adversely changed due to the stressor.

In short, recovery of affected populations from chemical stress may be rapid if the following conditions apply:

- The substance is not persistent in the aquatic environment, the exposure regime is short-term or pulsed, and the time between pulses is sufficient for recovery.
- The physicochemical environment and ecologically important food-web interactions are not altered by the stressor, or are quickly restored.
- The generation time of the populations affected is short.
- Delayed effects on reproduction due to short-term exposures can be excluded.
- There is a ready supply of propagules of eliminated populations through active immigration by mobile organisms or through passive immigration by, for example, wind and water transport.

From the evaluation above it follows that the most vulnerable aquatic species/populations are characterised by a low ability to avoid exposure in space and time, a high toxicological sensitivity and life cycle characteristics that hamper a fast recovery. When opting for the ETO, considerations for recovery are not necessary in the design and interpretation of higher tier studies, since there are no indications that within the same taxonomic group species with a low recovery potential show a higher

toxicological sensitivity than species with a high recovery potential (e.g. Brock et al., 2010c). When opting for ERO on the basis of micro-/mesocosm experiments, however, it needs to be critically evaluated whether representatives of potentially vulnerable populations are sufficiently covered in the micro-/mesocosm experiment. Further guidance on this is provided in chapter 9.

5.7. Implementation of the SPGs in this guidance document

In the tiered effect assessment scheme developed for this GD, in principle, all tiers are able to address the ETO (accepting negligible population-level effects) (section 5.5). However, the model ecosystem approach (micro-/mesocosm experiments) also allows to address the ERO when addressing risks to algae, vascular plants and invertebrates. In addition, appropriately designed and conducted micro-/mesocosm experiments are considered fit-for-purpose as a surrogate reference tier to calibrate the lower effect assessment tiers (on the basis of laboratory toxicity tests with standard and additional species). When using results of micro-/mesocosm tests to calibrate the lower tiers, or to derive an ETO-RAC, this will be done on the basis of negligible effects on the abundance and/or biomass of the most sensitive populations in the micro-/mesocosm test system. However, in this calibration, the extrapolation from surrogate reference tier (micro-/mesocosms) to the real reference (field) will need to be taken into account. In this GD, negligible effects in appropriately designed and conducted micro-/mesocosm experiments are equivalent to effect class 1 or effect class 2 responses for the most sensitive populations with the application of an appropriate AF for spatio-temporal extrapolation (see section 9.3.5.3 for more details).

When addressing the ERO in the SPGs for algae, vascular plants and invertebrates in edge-of-field surface waters, the indicated magnitudes and durations (temporal scale) of effects are characterised by interdependency (section 5.5). Since this interdependency may be influenced by factors like the vulnerability of the exposed organisms and a possible occurrence of multiple stressors, in this GD a prudent but pragmatic approach is followed when deriving a RAC from micro-/mesocosms on the basis of the ERO-RAC. It is proposed to base the ERO-RAC on an effect class 3A concentration (see section 9.3.3; i.e. a concentration that results in a maximum period of treatment-related effects < 8 weeks) for the most sensitive population-level or community-level endpoint by applying an appropriate AF to this concentration (see section 9.3.5.4 for more details). Note that, when applying the ERO, allowing a maximum period of treatment-related effects of less than 8 weeks (effect class 3) in particular will be a strict criterion for potentially sensitive univoltine invertebrates and rooted vascular plants. In addition, the application of the AF is meant to extrapolate the effect class 3A concentration in space and time so that other vulnerable populations (section 5.6) that may occur in the edge-of-field surface water at risk are sufficiently protected.

Although fungicides in particular may affect aquatic microbes it needs to be realised that tier 1 microbial test species are not mentioned in the revised data requirements (Commission Regulation (EU) No 283/2013). The implication of this is that it is assumed that by protecting populations of aquatic algae, vascular plants, invertebrates and vertebrates, the ecosystem services provided by bacteria and fungi are sufficiently protected. As there are no tier 1 data requirements for aquatic microbes available, no specific RA scheme is developed for them in this GD.

6. Exposure assessment

6.1. Introduction

PPP exposure assessment for the aquatic environment in the European Union is currently based on the FOCUS methodology (FOCUS, 2001). This is done for approval for a.s. at EU level. It is also used in some Member States for product authorisation, but also different exposure assessment procedures may be used. The FOCUS surface water methodology has not been reviewed by the PPR Panel of EFSA during the revision of the GD on Aquatic Ecotoxicology and the overall level of protection for approval of a.s. at EU level is therefore not clear. However, the methodology has been used in regulatory decision making throughout the last years and there is currently no alternative standardised exposure assessment methodology. Therefore, it is assumed that the $FOCUS_{sw}$ methodology will

continue to be used until updated or new methods become available and adopted by the SCFCAH and will replace the existing tools. FOCUS has a version control mechanism. When initiating assessments, the latest agreed exposure assessment tools, that is, latest versions of the FOCUS tools and documentation, should be applied.[25]

This chapter on exposure assessment in surface water is split into two sections, where the first one (6.2.1) gives a brief overview on how to perform the exposure assessment based on the current FOCUS surface water methodology and the second section (6.2.2) describes a guidance for the exposure assessment of metabolites that are formed in water and sediment. The intention of section 6.2 is only to give a short overview description of the FOCUS methodology. Thus, at the moment, the PPR Panel has in no way the ambition to evaluate the methodology nor does the PPR Panel have the intention to endorse the methodology. The PPR Panel, however, advises to critically evaluate and improve the surface water exposure assessment in the future.

The exposure assessment is often strongly influenced by the application method. FOCUS (2001) was developed mainly for spray applications. EFSA (2005b) provided additional guidance for using the FOCUS scenarios for non-spray applications.

6.2. FOCUS surface water scenarios and models

The FOCUS Surface Water Modelling Working Group defined a step-by-step procedure for the calculation of PECs in surface water (PEC_{sw}) (FOCUS, 2001). The procedure consists of four steps, whereby the first step represents a very simple approach using simple kinetics, and assuming a loading equivalent to a maximum annual application. The second step is the estimation of concentrations taking into account a sequence of loadings, and the third step focuses on more detailed modelling taking into account realistic 'worst-case' amounts entering surface water via relevant routes (run-off, spray drift, drainage). The third step considers substance loadings as foreseen in step 2, but it also takes into account the range of possible uses. The uses are, therefore, related to the specific and realistic combinations of cropping, soil, weather, field topography and aquatic bodies adjacent to fields. The fourth step accounts for risk mitigation measures. Notice that the FOCUS procedure is a stepped approach, not a tiered approach. The most important reason is that FOCUS (2001) has not proven that earlier steps are more conservative than later steps.

The aims of FOCUS (2001) for step 1 and 2 calculations were to represent 'worst-case loadings' and 'loadings based on sequential application patterns' respectively, but should not be specific to any climate, crop, topography or soil type. FOCUS (2001) considered the assumptions at both steps 1 and 2 as very conservative. Spray drift values are essentially based around drift numbers calculated from BBA (2000) and an estimation of the potential loading of PPPs to surface water via run-off, erosion and/or drainage. This loading represents any entry of PPP from the treated field to the associated water body at the edge of the field.

Step 3 requires the use of mechanistic models including PRZM, MACRO and TOXSWA.

Already at step 1 and 2, concentrations can be calculated, not only for the a.s., but also for metabolites formed in the soil before run-off/drainage occurs. The user must define the properties of the metabolite, including the maximum occurrence of the respective metabolite in soil studies and the ratio of the molecular masses of the parent and metabolite.

The fate of metabolites formed in the water body can also be taken into consideration at step 1 and 2. The formation will be calculated in a similar way based on the maximum occurrence of the metabolite in water/sediment studies.

[25] http://focus.jrc.ec.europa.eu/index.html

6.2.1. Description of the different steps

6.2.1.1. Step 1

The background of the FOCUS (2001) scenario properties on step 1 is a combination of existing concepts within the EU and Member States and measured datasets. A water depth of 30 cm overlying sediment of 5 cm depth with 5 % organic carbon (density: 0.8 g/cm³) is used; however, only 1 cm of the sediment is used in the calculations when calculating the distribution between water and sediment layer. When calculating PEC_{sed}, a depth of 5 cm is used, that is, a dilution of a factor 5 compared with the 1 cm used for the water sediment distribution. Sediment with 5 % organic carbon was selected in order to comply with existing RA approaches within the EU and existing ecotoxicity testing requirements for sediment-dwelling organisms.

Calculation of concentrations resulting from spray drift:

Spray drift deposition is expressed as the mass that enters the water per surface area of water, and assumed to be a certain fraction of the mass applied per surface area on the treated field.

Equation 1: $$C = \frac{D_{sd}\ App}{h}$$

where:

- App: application dose (rate) (g/m²)
- D_{sd}: spray drift deposition as fraction of the application dose, that is, mass deposited per surface area of surface water divided by mass deposited per surface area of field (-)
- h: surface water depth (m)
- C: concentration in surface water (g/m³)

Calculation of concentrations resulting from run-off, erosion or drainage:

Run-off, erosion or drainage loadings are defined as the PPP mass that enters the water and is expressed as a fraction of the total mass applied on the contributing treated field multiplied by the surface area of the contributing field:

Equation 2: $$C = \frac{App\ A_{field}\ L_{RO}}{h\ A_{sw}}$$

where:

- App: application dose (rate) (g/m²)
- L_{RO}: run-off loss as fraction of the applied PPP mass (-)
- A_{field}: area of the field contributing to run-off (m²)
- A_{sw}: surface area of surface water (m²)
- C: concentration in surface water (g/m³)

An explicit width or length of the water body for the initial step is not defined, because drift loadings are based upon a percentage of the application rate in the treated field. For run-off, erosion or drainage entries, only a fixed ratio A_{field}/A_{sw} of 10:1 is defined to reflect the proportion of a treated field from which PPPs are lost to surface water. This number was calibrated by model runs of PRZM, MACRO and TOXSWA (FOCUS, 2001).

At step 1, inputs of spray drift, run-off, erosion and/or drainage are evaluated as a single loading to the water body and 'worst-case' surface water and sediment concentrations are calculated. The loading to

surface water is based upon the number of applications multiplied by the maximum single-use rate except for compounds with a short half-life in sediment/water systems. If three times the degradation half-life ($3 \times DegT_{50}$) (combined water + sediment) is less than the time between individual applications, the maximum individual application rate is used to derive the maximum PEC as there is no potential for accumulation in the sediment/water system. For first order kinetics the value of $3 \times DegT_{50}$ is comparable to the $DegT_{90}$ value. Considering run-off loadings only run-off mass is entered into the stagnant 30 cm water, so no run-off water is added. This implies that exposure caused by run-off entries will be estimated in a conservative way by step 1.

Four crop groups (arable crops, vines, orchards and hops, representing different types of application technology) and aerial applications are separated into different drift classes for evaluation at step 1 and 2. Drift values have been calculated at the 90th percentile from BBA (2000) as summarised in Table 15 (FOCUS, 2001). The table indicates that no drift is assumed when the substance is incorporated or applied as granules or as a seed treatment. EFSA (2005b) concluded that dust drift may occur for such applications and provided computational procedures to estimate this route.

Table 15: Step 1 drift input into surface water based on the 90th percentile (from BBA, 2000)

Crop	Distance (m)[(a)]	Drift (%)[(b)]
Pome/stone fruit, early applications	3	29.2
Pome/stone fruit, late applications	3	15.7
Potatoes	1	2.8
Soybeans	1	2.8
Sugar beet	1	2.8
Sunflower	1	2.8
Tobacco	1	2.8
Vegetables, bulb	1	2.8
Vegetables, fruiting	1	2.8
Vegetables, leafy	1	2.8
Vegetables, root	1	2.8
Vines, early applications	3	2.7
Vines, late applications	3	8.0
Application, aerial	3	33.2
Application, hand (crop < 50 cm)	1	2.8
Application, hand (crop > 50 cm)	3	8.0
No drift (incorporation, granular or seed treatment)	1	0[(c)]

(a): Distance from edge-of-field to water body.
(b): Percentage of the application dose.
(c): 0 % drift for granular applications and seed treatments is not considered correct by EFSA (2005b).

In contrast, the run-off/erosion/drainage loading to the water body is set at 10 % of the application for all scenarios.

On the day of application, drift entries are assumed to be present only in the water phase in order to estimate a conservative peak concentration. One day later the compound is distributed between water and sediment.

In contrast, the run-off/erosion/drainage entry is distributed instantaneously between water and sediment at the time of loading according to the K_{oc} of the compound in order to simulate the process of deposition of eroded soil particles containing PPPs. In this way compounds are distributed directly between sediment and water. The relationship between K_{oc} and the distribution between water and sediment is calculated as follows:

Equation 3: Fraction of run-off in water $\dfrac{W}{(W + (S_{eff} \cdot bd \cdot oc \cdot K_{oc}))}$

where:

W:	mass of water (30 g)
S_{eff}:	mass of sediment available for partition (0.8 g)
oc:	mass fraction of organic carbon in sediment (0.05 g/g)
K_{oc}:	PPP organic carbon partition coefficient (cm^3/g)
bd:	bulk density of the sediment (g/cm^3)

6.2.1.2. Step 2

The surface water properties on step 2 are defined by FOCUS (2001) identically to step 1 so a static water body with a water depth of 30 cm, overlying sediment of 5 cm depth (density: 0.8 g/cm^3) with 5 % organic carbon. However, only 1 cm of sediment is used in the calculations when calculating the partitioning between water and sediment layer. When calculating PEC_{sed} a depth of 5 cm is used, that is, a dilution of a factor 5 compared with the 1 cm used for the water sediment distribution.

Also at step 2, the width of the water body is not defined because drift and run-off entries are calculated in a similar way based on a percentage of the application rate in the treated field. Also, the same ratio (10:1) is defined to reflect the proportion of a treated field from which PPPs are lost to surface water.

However, at step 2, inputs of spray drift, run-off, erosion and/or drainage are evaluated as a series of individual loadings comprising drift events (number, interval between applications and rates of application) followed by a loading representing a run-off, erosion and/or drainage event four days after the final application. Note that only run-off mass is entered into the stagnant 30 cm water, so no run-off water is added. This implies that peak exposure events caused by run-off entries will be estimated in a conservative way by step 2. Degradation is assumed to follow first-order kinetics in soil, surface water and sediment and the exposure assessor also has the option of using different degradation rates in surface water and sediment.

In order to prevent multiple worst-case assumptions for multiple application patterns, FOCUS (2001) defined different individual drift percentiles, dependent on the total number of applications per season, which, according to FOCUS (2001), represent the overall 90th percentile (see Table 16).

Table 16: Individual drift percentiles for multiple applications on step 2 (FOCUS, 2001)

Number of applications	Drift percentile[(a)]
1	90
2	82
3	77
4	74
5	72
6	70
7	69
> 8	67

(a): It is assumed that the individual drift events meet the overall 90th percentile.

As the procedure may result in lower predicted concentrations for multiple applications than for individual applications with the 90th drift percentile, the software automatically calculates both situations so that the user can select the higher value of the two.

Drift inputs are loaded into the water column where they are subsequently distributed between water and sediment according to the K_{oc} of the a.s.. However, the process of adsorption to sediment at step 2 is assumed to take longer than one day (as assumed at step 1). It is assumed that, following a drift event, the PPP is distributed in surface water into two theoretical compartments, 'available' for sorption to sediment and 'unavailable' for sorption to sediment.

Equation 4: $Ma_{sw} = K \cdot M_{sw}$

Equation 5: $Mu_{sw} = (1 - K) \cdot M_{sw}$

where:

M_{sw}: total mass of PPP in surface water (g),
Ma_{sw}: PPP mass available for sorption (g),
Mu_{sw}: PPP mass unavailable for sorption (g) and
K: fraction of PPP mass in water available for sorption (–)

K is set to two-thirds based on comparisons with laboratory sediment/water studies for weakly and strongly sorbing compounds (FOCUS 2001).

The partitioning between the sediment and the 'available' water compartment is calculated with the equation for the fraction in the run-off water given before using a mass of water of 30 g (so assuming a 30 cm water layer). This is not consistent with assuming that only two-thirds of the water is available for sorption as this would require a water depth of 20 cm. However, FOCUS (2001) considered the approach to be conservative for the PEC in surface water; for example, for a K_{oc} of 1 000 L/kg, the fraction remaining in the water layer after sorption equilibration is 0.43 for a water depth of 30 cm and 0.33 for a water depth of 20 cm.

The effect of the two-thirds available/one-third non-available water compartments for sorption is presented and discussed in more detail in Appendix B.

In contrast to step 1, the amount of PPP that enters the soil at step 2 is corrected for crop interception. For each crop, four interception classes have been defined depending on the crop stage. Crop interception will decrease the amount of PPP that reaches the soil surface and that thus ultimately enters the surface water body via run-off/drainage.

Four days after the final application, a run-off/erosion/drainage loading is added to the water body. This loading is a function of the residue remaining in the soil after all of the treatments (g/ha) and the region and season of application. The different run-off/drainage percentages applied at step 2 are listed in Table 17. They have been calibrated by FOCUS against the results of step 3 calculations as described in the FOCUS (2001) surface water report. For the calculation of the run-off event, Equation 4.2 is used, but instead of the application dose, the soil residue (in g/ha) is used which reflects the situation four days after application.

The user selects from two regions (Northern EU and Southern EU according to the definitions given for crop residue zones in the SANCO Document 7525/VI/95-rev.9, March 2011 (EC, 2011c)) and three seasons (March to May, June to September and October to February).

In common with step 1, the run-off/erosion/drainage entry is distributed between water and sediment at the time of loading according to the K_{oc} of the compound. An effective sorption depth of 1 cm is used for the distribution between both phases. In this way compounds of high K_{oc} are mostly added directly to the sediment whereas compounds of low K_{oc} are mostly added to the water column in the 'run-off/drainage' water. Contrary to spray drift entries, at run-off entries all mass in the water layer is available for sorption to sediment.

Table 17: Step 2 run-off/drainage input into surface water (from FOCUS, 2001)

Region/season	% of soil residue
North Europe, October–February	5
North Europe, March–May	2
North Europe, June–September	2

Region/season	% of soil residue
South Europe, October–February	4
South Europe, March –May	4
South Europe, June–September	3
No run-off/drainage	0

6.2.1.3. Step 3

For step 3, a selection of scenarios are defined based on a number of broad data sets that covered all areas of the European Community in 2001 (15 MS). According to FOCUS (2001), they should consider representative realistic worst-case situations and should take into account all relevant entry routes to a surface water body, as well as considering all appropriate target crops, surface water situations, topography, climate, soil type and agricultural management practices. However, due to the lack of comprehensive databases that characterise most of these agro-environmental parameters at the European level, when the scenarios were defined (1997–2001), they were not selected in a rigorous, statistically based manner. Instead a pragmatic approach was adopted, using very basic data sources together with expert judgement. All scenarios are represented by specific field sites for which monitoring data were available. Table 18 shows the inherent agro-environmental characteristics of the scenarios.

Table 18: Inherent agro-environmental characteristics of the surface water scenarios (from FOCUS surface water report (2001) table 3.2–6)

Scenario [(a)]	Meteoro-logical station	Mean spring and autumn temperature (°C)	Mean annual rainfall (mm)	Mean annual recharge (mm)	Slope (%)	Soil
D1	Lanna	< 6.6	600–800	100–200	0–0.5	Clay with shallow groundwater
D2	Brimstone	6.6–10	600–800	200–300	0.5–2	Clay over impermeable substrate
D3	Vredepeel	6.6–10	600–800	200–300	0–0.5	Sand with shallow groundwater
D4	Skousbo	6.6–10	600–800	100–200	0.5–2	Light loam over slowly permeable substrate
D5	La Jailliere	10–12.5	600–800	100–200	2–4	Medium loam with shallow groundwater
D6	Váyia, Thiva	> 12.5	600–800	200–300	0–0.5	Heavy loam with shallow groundwater
R1	Weiherbach	6.6–10	600–800	100–200	2–4	Light silt with small organic matter
R2	Valadares, Porto	10–12.5	> 1000	> 300	10–15	Organic-rich light loam
R3	Ozzano, Bologna	10–12.5	800–1000	> 300	4–10	Heavy loam with small organic matter
R4	Roujan	> 12.5	600–800	100–200	4–10	Medium loam with small organic matter

(a): D = Drainage, R= Run-off scenario.

Inputs to surface water bodies from spray drift are incorporated as an integral part of all of the scenarios based on the same tables as for the previous tiers (BBA, 2000). In addition to spray drift, the scenarios are characterised by either run-off/erosion (R) or drainage (D) entries.

For each location a maximum of two water body types is defined as shown in the following Table 19.

Table 19: Water bodies associated with scenarios (from FOCUS, 2001)

Scenario	Inputs	Water body type(s)[(a)]
D1	Drainage and drift	Ditch, stream
D2	Drainage and drift	Ditch, stream
D3	Drainage and drift	Ditch
D4	Drainage and drift	Pond, stream
D5	Drainage and drift	Pond, stream
D6	Drainage and drift	Ditch
R1	Run-off and drift	Pond, stream
R2	Run-off and drift	Stream
R3	Run-off and drift	Stream
R4	Run-off and drift	Stream

(a): All ditches and streams are assumed to have a length of 100 m, a width of 1 m and a variable, but minimum depth of 30 cm whereas the ponds are defined by surface water areas of 30 m × 30 m together with a depth of 100 cm.

For calculating substance entries into the surface water and for calculating time-dependent concentrations in the surface water bodies, different computer models are used. The currently recommended models (FOCUS, 2001) are MACRO for estimating the contribution of drainage, PRZM for the estimation of the contribution of run-off and erosion and TOXSWA for the estimation of the final PECs in surface waters and SWASH for the estimation of spray drift entries and as the overall user shell.

To facilitate the calculation of exposure concentrations at step 3 level, a software tool (SWASH) is available. It is an overall shell (user interface) combining all models involved in step 3 calculations. The main functions of the shell are:

- maintenance of a central PPP properties database;
- provision of an overview of all step 3 FOCUS runs required for use of a specific PPP on a specific crop;
- calculation of spray drift deposition onto various receiving water bodies; and
- preparation of input for the models MACRO (drainage entries), PRZM (run-off/erosion entries) and TOXSWA (fate in surface water).

Calculating drainage entries for TOXSWA with MACRO

MACRO is a general purpose leaching model that includes the effects of macropores (Jarvis, 1994, 2001). It was chosen by FOCUS to calculate drainage inputs to surface water bodies for the step 3 simulations because at that time it was the only FOCUS model that was able to simulate PPP losses through macropore flow. According to FOCUS this model was therefore suitable to cover the wide range of soil types included in the six drainage scenarios.

MACRO considers macropores as a separate flow domain assuming gravity flow of water and a simple power law function for the conductivity. Solute movement in the macropores is assumed to be dominated by mass flow, while the concentration of solutes in water entering the macropores at the soil surface is calculated using the 'mixing depth' concept, whereby the incoming rain perfectly mixes with the soil solution in a given depth of soil. MACRO describes the movement of water through the soil matrix using Richards' equation and solute transport with the convection–dispersion equation. Mass exchange between the flow domains is calculated using approximate first-order expressions

based on an effective diffusion path length. Sorption is described with a Freundlich isotherm, with the sorption sites partitioned between the two domains. Degradation is calculated using first-order kinetics.

Drainage from saturated soil layers is given as a sink term to the vertical one-dimensional flow equation using seepage potential theory (Leeds-Harrison et al., 1986) for saturated layers above drain depth and the second term of the Hooghoudt equation for layers below drain depth. Perched water tables are also considered. The bottom boundary condition utilised for the FOCUS surface water scenarios is a vertical seepage rate calculated as an empirical linear function of the height of the water table in the soil profile. PPP movement to the drains is calculated assuming perfect mixing in the lateral dimensions in each saturated soil layer.

FOCUS defined a 16-month assessment period for simulation of drainage inputs to surface waters. The weather data for the first 12 months of the assessment period were chosen by FOCUS to represent the 50th percentile year with respect to annual rainfall (the remaining four months were simply selected as the period following the selected 12-month period). As, especially for substances with high K_{oc} and $DegT_{50}$, the travel time of the PPP to the drains can be significantly longer than 16 months, FOCUS decided to employ a six-year warm-up period, in the same way as in the FOCUS groundwater scenarios (FOCUS, 2000). One of five different application methods can be selected by the user: ground spray, air-blast, granular, incorporated and aerial. Interception is assumed to be zero for both granular and incorporated applications. For air-blast applications and for ground and aerial sprays to perennial crops, the interception is assumed to always equal the maximum interception fraction. For annual crops, a fraction of the dose specified by the user is calculated as being intercepted by the crop canopy dependent on the application day(s) calculated by the pesticide application timer (PAT). This is given as a function of the method of application, a maximum interception reached at the maximum leaf area, and the leaf area index at the time of application.

Hourly values of water discharges through drains, and the PPP loads in the discharge during the assessment period are saved to an output file, which is then used as input to the surface water fate model TOXSWA.

Calculating run-off and erosion entries for TOXSWA with PRZM

The Pesticide Root Zone Model (PRZM) was selected to calculate run-off and erosion loadings for the the step 3 FOCUS surface water scenarios. It is a one-dimensional, dynamic, compartmental model that can be used to simulate chemical movement in unsaturated soil systems within and immediately below the root zone. It has two major components – hydrology and chemical transport. The hydrological component for calculating run-off and erosion is based on the US Department of Agriculture soil conservation service curve number methodology and a watershed-scale variation of the universal soil loss equation (USLE). Evapotranspiration is composed of evaporation from crop interception, evaporation from soil and transpiration from the crop. Water movement is simulated by the use of generalised soil parameters, including field capacity, wilting point and saturation water content (Carsel et al., 1995).

Hydrologic and hydraulic computations in PRZM are performed on a daily time step. To compensate for the long time step (compared with the other step 3 models) and to couple the run-off and erosion results simulated by PRZM with the transient hydrology incorporated in TOXSWA, the daily run-off and erosion time series output files are post-processed by FOCUS into a series of hourly run-off and erosion values by distributing the daily values linearly over a number of hours. This number equals the rainfall event size divided by an average rainfall intensity of 2 mm/h, (e.g. if there was 18 mm rainfall causing 4.1 mm run-off, the run-off event lasts 18 mm/2 mm per h = 9 hours and so, from midnight to 9 am there is 4.1 mm/9 h = 0.46 mm/h run-off).

The erosion loadings and chemical fluxes in run-off and erosion are handled in a similar manner.

The curve numbers used in PRZM were defined by FOCUS as a function of soil type, soil drainage properties, crop type and management practice. The curve numbers are used to determine a watershed retention parameter, which in turn determines the daily run-off as follows:

Equation 6: $S = 1000\ mm/RCN - 10\ mm$

where:

S: daily watershed retention parameter (mm)
RCN: run-off curve number (dimensionless, adjusted daily depending upon antecedent moisture).

Equation 7: $$Q = \frac{(P + SM - 0.2 \cdot S)^2}{P + SM + 0.8 \cdot S}$$

where:

Q: daily run-off (mm)
P: daily precipitation (mm)
SM: daily snow melt (mm)
S: daily watershed retention parameter (mm).

The following equation is used to calculate soil erosion by PRZM:

Equation 8: MUSS: $$Xe = 0.79 \cdot (Q \cdot q_p)^{0.65} A^{0.009} K \cdot LS \cdot SC \cdot P$$

where:

Xe: the event soil loss (metric tonnes/day)
Q: volume of daily run-off event (mm)
q_p: peak storm run-off (mm/h), determined from generic storm hydrograph
A: field size (ha)
K: soil erodability factor (dimensionless)
LS: length-slope factor (dimensionless)
SC: soil cover factor (dimensionless)
P: conservation practice factor (dimensionless).

This expression depends primarily upon daily run-off volumes and rates as well as the conventional USLE factors K, LS, C and p. The MUSS equation is a modification of the modified universal soil loss equation (MUSLE), developed by Williams (1975).

When calculating run-off and erosion losses, PRZM always runs over 20 years of data. Nevertheless, FOCUS selected one representative 12-month period for each use pattern being evaluated in step 3. The representative years selected for creation of PRZM output files for use by TOXSWA are given in the Table 20. For example, an application to maize, which occurs in June, would result in selection of the following 12-month period for scenario R3: June 1975 to June 1976.

Table 20: Selected years for creation of PRZM to TOXSWA (FOCUS 2001)

Scenario	Date of first application		
	March to May	June to September	October to February
R1	1984	1978	1978
R2	1977	1989	1977
R3	1980	1975	1980
R4	1984	1985	1979

According to FOCUS, the reason for not using the whole simulation period of 20 years as input was the computational requirements of TOXSWA.

The pesticide application timer (PAT) in MACRO and PRZM

When using MACRO, the user cannot enter explicit application dates. Instead the application has to be expressed as an application window which is used as input for PAT. The pesticide application timer eliminates a significant number of potential applications. According to FOCUS, this criteria in the PAT calculator results in selection of application dates which are the 60th to 70th percentile wettest days for non-irrigated crops and the 50th to 60th percentile wettest days for irrigated crops (the slightly lower percentile values for irrigated crops are the result of the additional number of wet days created by irrigation events for these crops).

When working with MACRO and PRZM, the user cannot enter application dates directly. Instead this is done by a similar PAT which uses an application window as input. The PAT then attempts to select appropriate application dates that meet two criteria:

- No more than 2 mm/day of precipitation should occur on any day within two days before or after an application;
- At least 10 mm of precipitation (cumulative) should occur within 10 days after an application.

If, however, no dates are found in the meteorological files that meet these criteria, the precipitation targets and timing in the two rules are progressively relaxed until acceptable application dates are found.

MACRO as well as PRZM calculate a fraction of the dose as being intercepted by the crop canopy. In both models the user can select the application methods of ground spray, air-blast, granular, incorporated and aerial. Interception is assumed zero for both granular and incorporated applications.

Calculating the fate of compounds in surface water with the TOXSWA model

The TOXSWA model describes the behaviour of PPPs in a water body at the edge-of-field scale, that is, a ditch, pond or stream adjacent to a single field. It calculates PPP concentrations in both the water and the sediment layers. In the water layer, the PPP concentration varies in the horizontal direction (varying in sequential compartments), but is assumed to be uniform throughout the depth and width of each compartment. In the sediment layer, the PPP concentration is a function of both horizontal and vertical directions.

TOXSWA considers four processes: (i) transport, (ii) transformation, (iii) sorption and (iv) volatilisation. In the water layer, PPPs are transported by advection and dispersion, while in the sediment, diffusion is included as well. The transformation rate covers the combined effects of hydrolysis, photolysis (in cases where this is accounted for in the experimental set-up used to derive this parameter value) and biodegradation, and it is a function of temperature. Sorption to suspended solids and to sediment is described by the Freundlich equation. Sorption to macrophytes is described by a linear sorption isotherm but this feature is not used in the TOXSWA model used for the FOCUS surface water scenarios. PPPs are transported across the water–sediment interface by advection (upwards or downwards seepage) and by diffusion. In the FOCUS surface water scenarios, transport across the water–sediment interface takes place by diffusion only.

The mass balance equations for the water and sediment layers are solved with the aid of a generalised finite-difference method. For the numerical solution, the water layer is divided into a number of nodes in the horizontal direction. Below each water layer node, an array of nodes is defined for the sediment layer. Distances between the nodes in the water and sediment layers are in the order of magnitude of metres and millimetres, respectively.

TOXSWA in FOCUS handles transient hydrology and PPP fluxes resulting from surface run-off, erosion and drainage as well as instantaneous entries via spray drift deposition. In order to simulate the flow dynamics in an edge-of-field water body in a realistic way, the field-scale system is defined as the downstream part of a small catchment basin.

The water body system in TOXSWA has been described with the aid of a water balance that accounts for all incoming and outgoing water fluxes. The incoming fluxes include the discharge from the upstream catchment basin (base flow component plus run-off or drainage component), the run-off or drainage fluxes from the neighbouring field. The outgoing fluxes are composed of the outgoing discharge of the water body and, if desired, a downwards seepage through the sediment. The water fluxes in the modelled system vary in time as well as in space, that is, with distance in the water body.

The water level in the water body varies in time, but it is assumed to be constant over the length of the water body. However, to prevent low water levels, a minimum water depth of 30 cm was defined for every stream and ditch scenario.

The TOXSWA model does not simulate the drainage or run-off/erosion processes itself, but uses the fluxes calculated by other models as entries into the water body system of TOXSWA. For this purpose, the PRZM in FOCUS model for run-off/erosion and the MACRO in FOCUS model for drainage create output files that list the water and mass fluxes as a function of time on an hourly basis. TOXSWA uses these output files as input to calculate the hydrological and PPP behaviour in the appropriate water body systems.

The variation of the water level in time has been calculated in two ways. For a pond, outflow is assumed to occur across a weir and the water level in the pond is derived with the aid of a classical Q(h) relation for a broad-crested weir (Ministère de la Coopération, 1984). In the case of a watercourse, the water level is calculated by a combination of uniform flow for which the Chézy–Manning equation can be applied in the upstream part of the watercourse and a backwater curve in front of a weir at the downstream end (Chow, 1959). The water levels in the FOCUS stream and ditch vary with time, but are assumed constant over their 100 m length (Adriaanse and Beltman, 2009).

The FOCUS ditch only occurs in FOCUS drainage scenarios where the land is relatively flat and, in most cases, relatively slowly drained. The ditch is assumed to be 100 m long and 1 m wide, with a rectangular cross-section. Its minimum depth is 0.3 m, implying that in all ditches an outflow weir maintains this minimum water level even during periods of very low discharge. It receives drainage fluxes from a 1 ha field adjacent to the ditch and from a 2 ha upstream catchment. PPP solute is only present in drainage waters from the 1 ha field adjacent to the ditch. The upstream catchment basin is assumed to be not treated with PPPs, therefore it is considered that only one third of the area considered in the ditch scenarios is treated with PPP.

The FOCUS stream occurs in the FOCUS drainage scenarios as well as the FOCUS run-off scenarios. Similar to the FOCUS ditch, the stream is assumed to be 100 m long and 1 m wide, with a rectangular cross-section. Its minimum depth is 0.3 m, implying that also in all streams a weir is located that maintains the 0.3 m water level even during periods of very low discharge. On one side of the stream a 1 ha field is located that delivers its drainage or run-off fluxes into the stream. This field is assumed to be treated with PPPs. The stream is also fed by the discharge of an upstream catchment basin of 100 ha which delivers its constant base flow plus variable drainage or run-off water fluxes to the stream. A surface area of 20 % of the upstream catchment basin is assumed to be treated with PPPs resulting in the dilution of edge-of-field drainage or run-off concentrations by an approximate factor of 5 before it enters the stream. The implications of PPP contribution from the upstream catchment is simplistically represented via an increase in the drift loading in the TOXSWA input file by a factor of 1.2 (e.g. additional 20 % loading).

Pond scenarios represent the simplest arrangement. Each 30 m × 30 m pond receives drainage or run-off waters with associated PPP in solution from a 4 500 m^2 contributing catchment. No PPP is present

in the base flow that enters the pond. For run-off scenarios, the pond also receives eroded sediment and associated PPP from a 20 m 'corridor' adjacent to the pond.

Dominance of entry routes

The FOCUS surface water scheme incorporates three potential routes of entry to surface water (spray drift, run-off and drainage). In aquatic exposure assessment using $FOCUS_{sw}$ modelling (and therefore RA), a.s. applied as sprays can generally be differentiated between those where spray drift is the dominant potential route of input to surface water and those where run-off and/or drainflow is the major potential input route.

If substances are applied as a spray and have a high potential for adsorption to soil particles (high K_{oc}), the spray drift route of input usually dominates. Profiles from step 3 $FOCUS_{sw}$ for these compounds give a distinct pulse of exposure for each spray drift event. The duration of this pulse is shortest for streams (dissipation driven primarily by advection) and longest for ponds (dissipation driven primarily by degradation and partitioning to sediment). For two reasons the initial PEC_{sw} values of a.s. in streams and ditches are higher than for ponds because of (1) the relative depth of the systems, and (2) the bigger surface water body area of the pond which results in lower spray drift deposition. This type of PEC_{sw} values is usually not sensitive to climate parameters. For the reasons relating to the method of selecting overall 90th percentile drift inputs for multiple application uses already discussed at step 2, where spray drift is the dominant entry route and there is relatively fast dissipation in the water body between entry events, simulations have to be carried out for a single application as well as multiple applications, to ensure that appropriate peak concentrations are generated and available for use in the RA. Unlike the step 2 tool, the SWASH shell does not generate these single application simulations automatically. The user has to define the single application simulations in addition to the multiple application simulations.

If, in contrast, the compound is characterised by high solubility in water, low K_{oc}, and relatively long $DegT_{50}$ in soil and the method of application favours run-off and drainage entries rather than spray drift (e.g. soil incorporation, low drift values as in field crops) run-off and/or drainflow becomes the major potential route of input to surface water. This type of PEC_{sw} value is sensitive to the rainfall pattern shortly after application as both processes are event driven.

6.2.1.4. Step 4

Step 4 simulations are usually performed considering the results of the FOCUS group on landscape and mitigation measures in ecological RA (FOCUS, 2007a, b).

Similar to the other tiers also for step 4, a software tool (SWAN) is recommended by FOCUS, which is available and developed on behalf of the European Crop Protection Association (ECPA). For interpretation of the mitigation of run-off in the FOCUS surface water scenarios as described by FOCUS in its landscape and mitigation report (FOCUS, 2007a; and also see Ter Horst et al., 2009). The software modifies the input and output files of the step 3 models TOXSWA and PRZM to consider drift and run-off buffer zones. The standard exposure reduction factors for run-off (water volume and PPP mass in run-off water) and erosion (eroded soil and PPP mass sorbed to eroded soil) as suggested by FOCUS (2007a, b) are shown in the Table 21.

Table 21: The 90th percentile worst-case values for reduction efficiencies for different widths of vegetated buffer strips and different phases of surface run-off (taken from FOCUS, 2007a)

	Buffer width (m)	
	10–12	**18–20**
Reduction in volume of run-off water (%)	60	80
Reduction in mass of PPP transported in aqueous phase (%)	60	80
Reduction in mass of eroded sediment (%)	85	95

	Buffer width (m)	
	10–12	**18–20**
Reduction in mass of PPP transported in sediment phase (%)	85	95

It should be noted that whilst SWAN can be used to parameterise run-off loadings with greater reduction values than those indicated in Table 21, FOCUS (2007a) prescribes a ceiling on run-off mitigation of 90 % run-off entry reduction. Regulatory practice means that water volume and substance solute mass should not be reduced by > 80 % and mass of eroded sediment/substance mass in the eroded sediment should not be reduced by > 95 %. Alternatively, in regulatory practice, it has been accepted to have all four of these parameters reduced by a maximum of 90 %. In regulatory practice, other combinations of reduction approaches might also be accepted, but it is the responsibility of the applicant to clearly demonstrate that the approach that they have taken, respects the ceiling of not reducing run-off inputs by more than 90 % of that calculated by PRZM at step 3.

SWAN can also handle drift reduction due to the use of more advanced nozzle techniques (low-drift nozzles). In addition to the entry routes considered in the first three steps, the exposure via air for volatile substances, using the recommendations of the FOCUS air group (FOCUS, 2008), is current practice.

The effect of drift buffer zones (i.e. no-spray buffer zones) can be considered in SWAN for distances up to 100 m from the surface water body. The model considers the same reduction rates as in the FOCUS SWASH tool both based on BBA (2000). It should be noted that whilst SWAN can be used to parameterise drift buffer zones up to 100 m and the effects of low drift nozzles can be combined with drift buffer zones to reduce spray drift inputs still further, FOCUS (2007a) prescribes a ceiling on spray drift mitigation. This prescription is that spray drift cannot be mitigated such that the mass per unit area reaching the water body surface is < 5 % of this value that would be calculated for the crop of interest using the FOCUS defined baseline distance for that crop (1–6 m, i.e. the ceiling for spray drift mitigation is 95 %).

As at step 3, when a use pattern includes multiple applications, it can also be necessary to simulate a single application as well as multiple applications at step 4, to ensure that appropriate peak concentrations are generated and available for use in the RA. The need for this procedure to be necessary reduces as the extent of spray drift mitigation implemented increases.

6.2.2. Assessment of metabolites by FOCUS surface water modelling

Already, at step 1 and 2, concentrations can be calculated, not only for the a.s., but also for metabolites formed in the soil before run-off/drainage occurs. The user must define the properties of the metabolite, including the maximum occurrence of the respective metabolite in soil studies and the ratio of the molecular masses of parent and metabolite.

The fate of metabolites formed in the water body can also be taken into consideration at step 1 and 2. The formation will be calculated in a similar way based on the maximum occurrence of the metabolite in water/sediment studies.

Based on this information, equivalent rates for metabolites formed in soil (Equation 9) or water/sediment (Equation 10) are calculated which are then used to calculate concentrations in surface water and sediment for metabolites.

Equation 9: $$App_{met,runoff/drainage} = \frac{App\,M_{met}\;R_{max,soil}}{M_{par}}$$

where:

$App_{met,runoff/drainage}$: equivalent dose (rate) of metabolite for run-off (g/m²)
App: parent application dose (rate) (g/m²)
M_{par}: molecular mass of the parent (g/mol)
M_{met}: molecular mass of the metabolite (g/mol)
$R_{max,soil}$: maximum residue fraction of the metabolite observed in soil studies (-)

Equation 10: $$App_{met,drift} = \frac{App\ M_{met}\ R_{max,\ wholesystem}}{M_{par}}$$

where:

$App_{met,drift}$: equivalent dose (rate) of metabolite for drift (g/m²)
App: parent application dose (rate) (g/m²)
M_{par}: molecular mass of the parent (g/mol)
M_{met}: molecular mass of the metabolite (g/mol)
$R_{max,whole\ system}$: maximum residue fraction (whole system) of the metabolite observed in water/sediment studies (-)

For the calculation of surface water concentrations, in principle, the same equations are used for metabolites as described for parent compounds earlier in this document.

In the current software, drift inputs are considered for water metabolites and run-off inputs for soil metabolites. That means, for water metabolites the current step 1 and 2 software, calculator version does not consider run-off entries from parents.

If all the necessary fate and behaviour properties of a metabolite needed for the step 1 & 2 software calculator are not available, conservative default values can be accepted, should these allow a satisfactory risk characterisation to be completed. This can be possible when the necessary effects data of appropriate quality are available for the metabolite and these indicate low hazard. Default values that have been agreed by Member State competent authorities to be used as input in step 1 & 2 software calculations since 2002, are a K_{oc} of 10 L/kg with respect to water column calculations and 10 000 L/kg with respect to sediment calculations. Should metabolite $DegT_{50}$ values be unavailable, a default value of 300 days has been used as input in step 1 & 2 software.

6.2.2.1. Metabolites at step 3 and step 4

The MACRO model can deal with one parent compound and one metabolite in one single simulation sequence. If more than one metabolite is being formed then another simulation sequence should be performed, for the same parent compound, but the second metabolite. It prepares an output file, listing the metabolite drainage fluxes as a function of time that TOXSWA reads in. Secondary metabolites cannot be considered directly in MACRO. Instead they have to be considered like primary transformation products.

The PRZM model can handle two metabolites simultaneously. Either two metabolites are formed from the parent compound, or the first metabolite degrades into a second metabolite. In both cases, PRZM prepares two separate output files that list the metabolite run-off fluxes as a function of time. The TOXSWA model can read these files and thus account for the fate of metabolites in the water body.

A parent compound that is deposited on the surface area of the FOCUS water body dissolves into the water and metabolites are formed. Additionally, metabolites may enter the water body via various entry routes. The currently released version of the TOXSWA model, FOCUS_TOXSWA_3.3.1, can only handle metabolites formed in soil that entered the water body via drainage or runoff/erosion.

The TOXSWA_M model (FOCUS successor version after 3.3.1) can handle a large number of metabolites in one run; these metabolites may be formed in soil and enter the water body via drainage

or runoff and erosion or they may be formed in the water and/or sediment. Metabolites formed in the water body may be formed in so-called consecutive reactions, as well as in so-called parallel reactions. The user needs to set up the list of compounds to consider and indicate, in matrix form, which couples of compounds are connected by a reaction. Additional inputs needed are the molar formation fractions and the physicochemical properties of the metabolites. All present output for the parent is also available for the metabolites, that is, concentration as a function of time in water (total and in dissolved phase), sorbed to suspended solids, and in sediment (in pore water or sorbed to sediment mass), mass present in different phases in water and sediment as a function of time, and full mass balances for water.

There is currently no harmonised procedure regarding how to consider metabolites at step 4. FOCUS (2001) provided advice on how to resolve this problem. However, the procedures are complicated and are infrequently used in current exposure assessment. Instead, different methodologies are followed at the present time.

7. Data requirement for active substances and formulations and tier 1 effect assessment

7.1. Introduction to data requirements as laid down in Commission Regulations (EU) 283/2013 and 284/2013 for approval of active substances and plant protection products and related OECD guidelines

The data requirements state that the ecotoxicological RA shall be based on the risk that the a.s. and the formulated PPP poses to non-target organisms. In carrying out a RA it is necessary to compare toxicity with exposure.

The specific data requirements for Regulation (EC) No 1107/2009 concerning the placing of PPPs on the market are laid down in Commission Regulation (EU) 283/2013 for the dossier to be submitted for the approval of a.s. contained in PPPs and in the Commission Regulation (EU) 284/2013 for the authorisation of PPPs.

7.1.1. Test guidelines

Studies should always be assessed according to recognised test guidelines. A list of test methods and GDs that are relevant is provided in two Commission Communications on the implementation of Regulation 283/2013[26] and 284/2013[27]. Specific guidance for testing difficult substances and mixtures in aquatic toxicity tests is provided in the OECD guidance document 23 (OECD, 2000).

An overview of the obligatory and additional toxicity tests (that should be provided under certain circumstances) by the applicant is presented in Table 22.

[26] Commission Communication in the framework of the implementation of Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ C 95, 3.4.2013, pp. 1–20.

[27] Commission communication in the framework of the implementation of Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ C 95, 3.4.2013, pp. 21–37.

Table 22: Ecotoxicity studies required for active substances under certain circumstances (for metabolites see section 10.2).

	Acute toxicity test to fish (rainbow trout)	Acute toxicity test to *Daphnia*	Toxicity test to green alga	Acute toxicity test to additional arthropod, e.g. *Chironomus* ssp. or *Americamysis bahia*	Toxicity test to algae (not green alga, e.g. diatom *Navicula pelliculosa*)	Toxicity test to *Lemna*	Toxicity on other macrophyte species (e.g. *Myriophyllum* or *Glyceria*)	Fish early-life stage toxicity test (ELS)	Long-term/chronic toxicity test on *Daphnia magna* or in the case of two tested arthropod species in the acute situation, a test with the species that showed the lowest endpoint	Chronic spiked toxicity test on *Chironomus riparius* or *Lumbriculus* spp	Fish short-term reproductive assay, 21-d fish assay or fish sexual development test	Fish full life cycle toxicity test (FFLC)
Every substance	x	x	x									
Substances with insecticidal mode of action (MOA)				x[(a)]								
Substances with a herbicidal MOA or plant growth regulators					x	x						
Substances with a herbicidal MOA for which *Lemna* is not sensitive or there is expected uptake by the roots of submerged macrophytes[(b)]							x[(b)]					
Where exposure of surface water is possible and the substance does not hydrolyse instantly ($DegT_{90}$> 1 d)								x[(c)]	x			
Accumulation of the substance in sediment indicated or predicted from fate studies[(d)]										x[(e)]		
Substances which are suspected to interfere with the moulting hormones (e.g. insect growth regulator)										x[(f)]		
Substances identified as an endocrine active substance (EAS) where the known MOA may be expected to impact fish sexual											x	

	Acute toxicity test to fish (rainbow trout)	Acute toxicity test to *Daphnia*	Toxicity test to green alga	Acute toxicity test to additional arthropod, e.g. *Chironomus* ssp. or *Americamysis bahia*	Toxicity test to algae (not green alga, e.g. diatom *Navicula pelliculosa*)	Toxicity test to *Lemna*	Toxicity on other macrophyte species (e.g. *Myriophyllum* or *Glyceria*)	Fish early-life stage toxicity test (ELS)	Long-term/chronic toxicity test on *Daphnia magna* or in the case of two tested arthropod species in the acute situation, a test with the species that showed the lowest endpoint	Chronic spiked toxicity test on *Chironomus riparius* or *Lumbriculus* spp	Fish short-term reproductive assay, 21-d fish assay or fish sexual development test	Fish full life cycle toxicity test (FFLC)
development and/or reproduction												
Substances exhibiting endocrine activity in 21-d fish screening assay or fish sexual development test (see above)												x

(a): The PPR Panel recommends the use of *Chironomus* for testing compounds with insecticidal mode of action, if data on *Americamysis bahia* are not already available.

(b): Additional testing may be required by the national competent authorities on other macrophyte species depending on the mode of action of the substance, or if clear indications of higher toxicity are apparent to dicots (e.g. auxin inhibitors, broad leaf herbicides), or other monocots (e.g. grass herbicides) plant species from efficacy or non-target terrestrial plant testing. Additional aquatic macrophyte testing may be undertaken on dicots (e.g. *Myriophyllum spicatum* or *M. aquaticum*) or monocots (e.g. *Glyceria maxima*) as appropriate. The need to perform such studies shall be discussed with national competent authorities.

(c): Unless a fish full life cycle (FFLC) test is provided. A FFLC test may be required depending upon the persistence and bioaccumulative potential of the substance. The Panel recommends that FFLC tests may be required where the BCF is > 1 000, the elimination during the 14-d depuration phase in the bioconcentration study is < 95 % or the substance is stable in water or sediment ($DegT_{90}$ > 100 d). Long-term exposure may also occur for substances which show degradation in water and sediment if leaching from drainpipes contributes significantly to the exposure in surface water. So if long-term exposure is expected based on the predicted field exposure profile, an FFLC study might be required as well. However, it is not yet possible to provide rules of thumb for the significance of leaching from drainpipes based on the $DegT_{50}$ in soil, the K_{om} and other relevant substance characteristics. Development of such rules of thumb may be helpful for the RA.

(d): Water/sediment study showed > 10 % of applied radioactivity at or after day 14 present in the sediment and chronic daphnia test (or other comparable study with insects) EC_{10} (or NOEC) < 0.1 mg/L. For the time being, the guidance as given in the former SANCO guidance (2002) should be followed. This might be revised in the future, depending on the PPR Panel opinion on sediment effect assessment under development (EFSA-Q-2012–00959).

(e): The PPR Panel recommends *Chironomus* for compounds with insecticidal activity and *Lumbriculus* for a.s. with fungicidal activity (based on data presented in Maltby et al., 2005, 2009).

(f): If an a.s. is an insect growth regulator, data should be preferably provided for *Chironomus*.

7.2. Standard toxicity tests with aquatic organisms

The following toxicity tests should be submitted for every substance even when it is not expected that preparations containing it could reach surface water following the proposed conditions of use: acute toxicity to fish (*Oncorhynchus mykiss*), acute toxicity for *Daphnia* species (preferably for *Daphnia magna*) and effects on the growth for a green alga (e.g. *Pseudokirchneriella subcapitata*).

A limit test at 100 mg substance/L may be performed when the results of a range finding test indicate that no effects are to be expected. In order to minimise fish testing, a threshold approach to an acute fish test should be considered (OECD, 2010). An acute fish limit test should be conducted at 100 mg substance/L or at an appropriate concentration selected from aquatic endpoints following consideration of the threshold exposure. Other considerations for setting the limit in a limit test could be compound properties (e.g. water solubility (see also sections 11.2 and 11.4)), or needs for RA. When mortality is detected in the fish limit test, an acute fish dose–response toxicity study should be performed to determine an LC_{50} for use in RA.

The endpoints required for toxicity tests in the 'revised data requirements' for the standard tests and the other requested additional types of tests are EC_{10}, EC_{20} and EC_{50}, together with their 95 % confidence intervals (or an explanation if they cannot be estimated) and corresponding NOEC values. How these should be determined is explained below.

7.2.1. Reasoning for the introduction of new endpoints (EC_x)

Traditionally, the responses measured in chronic laboratory ecotoxicity tests have been expressed as the No Observed Effect Concentration (NOEC) and/or the lowest observed effect concentration (LOEC). The PPR Panel recommends that, based on laboratory toxicity tests, RA for all groups of organisms should use the EC_x approach in preference to the NOEC/LOEC approach (EFSA, 2009a). The recommendations for the choice of effect percentile in the EC_x for each group of organisms should be based on an analysis of existing study data for each of these groups and should take into consideration the following issues:

- The percentile (x in EC_x) for each group of organisms should be chosen so as to ensure a level of protection consistent with the aims of the regulations, taking into account the conservatism of other parts of the RA.
- The choice of EC_x should take into account the reliability of the estimates that can be provided by standard test methods.
- The procedure for using the chosen EC_x in the RA should take account of the quality of the estimates available for each substance, for example by examining confidence intervals for the EC_x and possibly using these in the RA.
- As existing study methods were not designed to estimate EC_x, it is expected that a proportion of existing studies will not provide a usable (or even any) estimate for the preferred EC_x. For reasons of cost and animal welfare, the procedure should be designed to minimise or, if possible, prevent any need for retesting in such cases. For example, the procedure could include provisions to use alternative EC_x or even the NOEC in such cases, together with appropriate adjustments to the RA (e.g. different AF) to provide the required level of protection.
- The desirability of harmonising with EC_x approaches used under other EU legislation (e.g. REACH; EC, 2006) should be considered, unless there are good reasons to differ.

In order to provide the flexibility to accommodate these issues in the final procedures and for the Commission Regulation (EU) No 546/2011 to be revised, the choice of EC_x, and whether or not to use confidence intervals on the EC_x should ideally be left open until an appropriate analysis has been performed, on which recommendations can be based. The possibility of using the NOEC where necessary should also be retained. This will have the added benefits of (a) providing an indication of the slope of the dose–response curve (by comparing the EC_{10}, EC_{20} and EC_{50}) and (b) helping the transition from NOEC to EC_x by presenting both together.

For practicality reasons, it is proposed to use EC_{10} in the chronic RA scheme (except for plants) for the time being, or when the EC_{10} is not available, the NOEC in accordance with the Technical Guidance Document 27 to the Water Framework Directive 2000/60/EC (EC, 2011b) until new knowledge on the choice of EC_x becomes available. The same AF is used to derive a RAC from an EC_{10} or a NOEC.

7.2.2. Ionisable substances

Ionisable substances are substances that possess either weak acidic or basic functional groups and depending on pH can be present in both its dissociated and non-dissociated form. The dissociated and non-dissociated species may have different water solubilities and partition coefficients, and therefore bioavailability and toxicity. It may be possible to identify ionisable groups within the structure of the molecule. Groups containing, for example, oxygen, sulphur, phosphor or nitrogen atoms such as thiol (SH), sulphonate (SO_3H), hydroxyl (OH), carboxyl (COOH) or amine (NH_2) groups, are all potentially ionisable.

Substances that ionise at naturally relevant pH values can vary by several orders of magnitude in bioconcentration and toxicity. Therefore, it is essential to know or estimate the pK_a to evaluate the degree of ionisation in surface waters. The extent of ionisation may vary according to pH or the level of counter ions in the media, and relatively small changes may significantly alter the equilibrium between dissociated and non-dissociated species.

As unionised organic forms tend to be more hydrophobic than the ionised forms, the solubility and bioavailability of the substance may vary dramatically even between environmental extremes in pH. Consideration should be given to appropriate pHs (to be) used in the test as solubility may be lower but toxicity may be higher in the unionised form than in the ionised form. Therefore, testing of bioconcentration and toxicity of ionisable substances should preferably be conducted at a pH consistent with the more toxic form of the substance whilst remaining within the range required to maintain the health of the control organisms. One should keep in mind that sometimes uptake is influenced, so testing under conditions of the more toxic form does not necessarily lead to the higher effects. A stable pH is important to ensure that the balance between dissociated and non-dissociated forms of the substance is maintained.

7.2.3. Fish

7.2.3.1. Acute toxicity to fish

A 96-h test on rainbow trout (*Oncorhynchus mykiss*) with the a.s. shall always be carried out, even when it is not expected that the compound will end up in the surface water. In that case it will be used for classification and labelling. Consideration should be given to allow the reduction of animal testing (see also section 11.4).

7.2.3.2. Chronic toxicity to fish

Circumstances in which required

A long-term or chronic toxicity study on fish is required for all a.s. where exposure of surface water is likely and the substance is deemed to be stable in water, that is, there is less than 90 % loss of the original substance over 24 h via hydrolysis at all relevant pH values (hydrolysis $DegT_{90}$ > 24 h). An early-life stage study is always required in these circumstances; however, if a full fish life cycle study has been generated, an early-life stage study is not required.

The fish early-life stage test should determine effects on development, growth, survival and behaviour, and details of observed effects on fish early-life stages.

A full fish life-cycle test may be required depending upon the persistence and bioaccumulative potential of the substance.[28]

[28] The Panel recommends that FFLC-tests may be required where the BCF is > 1000, the elimination during the 14 day depuration phase in the bioconcentration study is < 95 % or the substance is stable in water or sediment ($DegT_{90}$ > 100 days). Long-term exposure may also occur for substances which show degradation in water and sediment if leaching from drainpipes contributes significantly to the exposure in surface water. So if long-term exposure is expected based on the predicted field exposure profile, a FFLC study might be required as well. However, it is not yet possible to provide rules of

For substances that fulfil the screening criteria on either of the fish screening assays, or where there are other indications of endocrine disruption (see paragraph 5.6), appropriate additional endpoints should be included in the test as recommended by the OECD conceptual framework in support of testing and assessment of potential endocrine disrupters should be discussed with the national competent authorities. The test conditions should be designed to reflect concerns identified in already available aquatic toxicity tests, mammalian and bird toxicology studies and other information. The exposure should be either continuous or, when appropriate arguments are provided, pulsed exposure considering the sensitive life stages according to the mode of action.

7.2.4. Amphibians

Even if the revised data requirements (Commission Regulation (EU) 283/2013) do not request toxicity tests for amphibian species, amphibians should be included in the aquatic and terrestrial RA of PPPs. Assessment of the risk to amphibians should be based on any existing relevant information. Available relevant data, including data from the open literature, for the substance under consideration should be presented and taken into account in the RA according to EFSA (2011). Terrestrial life stages of amphibians will be addressed in a future GD on PPP RA for amphibians and reptiles (EFSA-Q-2011–00987) under the mandate of the revision of the GD on terrestrial ecotoxicology. In this guidance, only aquatic life stages are addressed.

An analysis of acute toxicity data for a large number of amphibian species (Fryday and Thompson, 2012) and comparison with fish acute toxicity data (see Appendix C) shows that the rainbow trout is a good surrogate test species for predicting the acute toxicity of PPPs for larval stages of amphibian species living in the aquatic compartment of the environment. Similar results were found by Aldrich (2009). By using the same AFs as have been applied for fish, the achieved level of protection will be the same for both groups of organisms. The assessment is only valid for acute toxicity (mortality) and will not necessarily be predictive of chronic toxicity. However, a recent study indicates that the same is also applicable for chronic toxicity (Weltje et al., 2013).

If a refinement of the RA needs to be performed for fish, it needs to be investigated whether the refined RA for fish still covers amphibians. For amphibians, refined RA could start with carrying out a toxicity test for an amphibian species, or the same approach could be adopted as is relevant for fish higher tier assessment. More guidance is given related to the use of the SSD approach (section 8.4.5) and for refined exposure studies (section 9.2).

7.2.5. Aquatic invertebrates

A 48-h acute test with *Daphnia magna* (e.g. OECD guideline 202) always has to be carried out, even when it is not expected that the compound will end up in surface water. In that case it will be used for classification and labelling.

In addition, for a.s. with an insecticidal mode of action or which show insecticidal activity[29] a second arthropod species shall be tested, for example chironomid larvae (OECD guideline 235) or mysid shrimps (*Americamysis bahia*). The PPR Panel recommends that, preferably, a chironomid test for a.s. with insecticidal activity should be used. The PPR Panel proposes that if *Daphnia* is an order of

thumb for the significance of leaching from drainpipes based on the $DegT_{50}$ in soil, the K_{om} and other relevant substance charachteristics. Development of such rules of thumb may be helpful for the RA.

[29] Data for the non-target arthropods could be used for assessing the potential insecticidal activity of a compound. For most of the compounds the two standard non-target arthropods are tested (*Typhlodromus pyri* and *Aphidius rhopalosiphi*). When the quotient of the application rate multiplied by a multiple application factor (MAF) and the LR_{50} is greater than 2 the compound could be considered as having insecticidal activity. In addition efficacy studies with other insects or studies carried out with insects in the screening process could be another source for assessing potential insecticidal activity. In addition data for bees can also be used, if the hazard quotient (HQ) of the application rate in g/ha divided by the acute toxicity (LD_{50}) ≥ 50 the compound can be considered to have insecticidal activity.

magnitude more sensitive than algae or fish, a second arthropod species is required as well, since this might also indicate insecticidal acitivity.

A long-term or chronic toxicity study on aquatic invertebrates should be provided for all a.s. where exposure of surface water is likely and the substance is deemed to be stable in water, that is to say there is less than 90 % loss of the original substance over 24 h via hydrolysis (hydrolysis $DegT_{90}$> 24 h, see point 7.2.1.1 of the data requirements).

A chronic toxicity study should be submitted on one aquatic invertebrate species. If acute tests have been conducted on two aquatic invertebrate species (e.g. *Daphnia* and *Chironomus*) the acute endpoints should be taken into account (see point 8.2.4 of the data requirements) in order to determine the appropriate species to be tested in the chronic toxicity study. The PPR Panel proposes to select the more sensitive species in case the difference in acute toxicity is more than a factor of 10. If a chronic test with *Chironomus* is selected, the exposure regime should be long-term and the EC_{10}/NOEC expressed in terms of mean exposure concentrations during the test (i.e. as in OECD guideline 233 with quartz sand substrate and chronic exposure).

If the test substance is suspected of interfering with moulting hormones, that is, it is an insect growth regulator, or if the test substance has other effects on insect growth and development, an additional study on chronic toxicity shall be carried out using relevant non-crustacean species such as *Chironomus* spp.

7.2.5.1. Toxicity studies with sediment-dwelling organisms

When accumulation of an a.s. in aquatic sediment is indicated or predicted by environmental fate studies[30], the impact on a sediment-dwelling organism shall be assessed. The chronic risk to *Chironomus riparius* (OECD 218, 219) or *Lumbriculus* spp. (OECD 225) shall be determined. PPR recommends *Chironomus* for compounds with insecticidal activity and *Lumbriculus* for a.s. with fungicidal activity (based on information provided by Maltby et al., 2005, 2009).

An appropriate alternative test species may be used where a recognised guideline is available. The a.s. should be applied to either the water or the sediment phase of a water/sediment system and the test should take account of the major route of exposure (i.e. tests could be either water spiked or sediment spiked).

The key endpoint from the study should be presented in terms of mg substance/kg dry sediment and mg substance/L water.

This GD focuses on exposure via the water phase. A scientific opinion addressing the effect assessment for sediment organisms in detail will be developed by the PPR Panel in the near future.

There are two OECD guidelines available for testing either spiked water or spiked sediment (OECD 218 and 219). First, instar chironomid larvae are exposed. Chironomid emergence and development rate is measured at the end of the test. The maximum exposure duration is 28 d for *C. riparius* and *C. yoshimatsui*, and 65 d for *C. dilutus* (formerly *C. tentans*).

An OECD draft guideline is available for assessing the effects of life-long exposure of chemicals to the freshwater dipteran *Chironomus* spp., fully covering the first generation and the early part of the second generation (OECD guideline 233).

[30] As described in Table 22, water/sediment study showed > 10 % of applied radioactivity at or after day 14 present in the sediment and chronic daphnia test (or other comparable study with insects) NOEC< 0.1 mg/L. For the time being, the guidance as given in the former SANCO guidance (EC, 2002a) should be followed. This might be revised in the future, depending on the PPR Panel opinion on sediment effect assessment under development (EFSA-Q-2012-00959).

A sediment-spiked test for *Lumbriculus variegatus* is described in OECD guideline 225. *Lumbriculus* spp. are exposed for 28 d and effects on reproduction and biomass are observed.

7.2.6. Standard toxicity tests with algae

A 72- to 96-h test should always be carried out on one green alga (such as *Pseudokirchneriella subcapitata*, synonym *Selenastrum capricornutum*). For a.s. that exhibit herbicidal activity[31], a test on a second species from a different taxonomic group should be performed, such as a diatom (e.g. *Navicula pelliculosa*).

The algal test (OECD guideline 201) is a short-term test, although it provides chronic endpoints. The preferred observational endpoint in this study is algal growth rate inhibition because it is not dependent on the test design, whereas an endpoint based on biomass is dependent on both the growth rate of the test species as well as test duration and other elements of test design. Often, both growth rate EC_{50} (E_rC_{50}) and biomass (E_bC_{50}) endpoints are reported, however, the latter should not be used. The reason is that direct use of the biomass concentration without logarithmic transformation cannot be applied to an analysis of results from a system in exponential growth (ECHA, 2008). Where only the E_bC_{50} is reported, but primary data are available, a reanalysis of the data should therefore be carried out to determine the E_rC_{50}. However, if only E_bC_{50} values are presented, this value can also be used in the RA. The result for the endpoint biomass (growth) is generally somewhat lower than the growth rate and can therefore be considered as a conservative value. The OECD guideline 201 now provides methods for calculating growth rate and yield. Growth rate is the preferred endpoint to be used, yield is only included for cases where specific regulatory requirements in some countries may need to be fulfilled.

7.2.7. Standard toxicity tests with macrophytes

If a substance is a herbicide, a plant growth regulator or if it shows herbicidal activity[31], a test with *Lemna* spp. should be carried out. Tests could be performed according to OECD test guideline 221. In this test, exponentially growing plant cultures of the genus *Lemna* (*Lemna gibba* and *Lemna minor* usually) are allowed to grow as monocultures over a period of seven days. The objective of the test is to quantify substance-related effects on vegetative growth over this period based on assessments of selected measurement variables. This study includes the counting of the frond number and measurement of at least one other variable (total frond area, dry weight or fresh weight) using the lowest of these endpoints for the RA. Growth rate is the preferred endpoint to be used since it is more robust considering varying test conditions. It should be calculated (as E_rC_{50}) on the basis of the most sensitive endpoint. Yield is only included in the OECD guideline 221 for cases where specific regulatory requirements in some countries may need to be fulfilled.

According to the data requirements, additional testing may be required by the Member State competent authorities on other macrophyte species depending on the mode of action of the substance, or if clear indications of higher toxicity are apparent to dicotyledonous (for example auxin inhibitor, broad leaf herbicides) or other monocotyledonous (e.g. grass herbicides) plant species from efficacy or testing with terrestrial non-target plants. Additional aquatic macrophyte species tests may be undertaken on a dicotyledonous species, such as *Myriophyllum spicatum, Myriophyllum aquaticum* or a monocotyledonous species, such as aquatic grass *Glyceria maxima,* as appropriate.

An analysis using SSDs by Giddings et al. (2013) indicated that neither *Lemna gibba* nor *Myriophyllum spicatum* is consistently among the most sensitive macrophyte species for all the herbicides and fungicides included in their analysis. The *L. gibba* EC_{50} is within a factor of 10 of the HC_5 of the macrophyte SSD for 6 of 11 chemicals investigated, indicating that RA based on *Lemna*

[31] If in one or more of the screening or efficacy tests with vascular plant species showing ≥ 50% phytotoxic effects at the maximum recommended application rate (MRR) or higher, the a.s. is considered to have herbicidal activity (see EPPO scheme for higher plants (EPPO, 2003)).

might not always be protective for other macrophytes. The *M. spicatum* EC_{50} was within a factor of 10 of the macrophyte HC_5 for 5 of 11 chemicals.

The PPR Panel of EFSA recommends to use *Lemna* sp. as the default macrophyte test species and to follow the recommendations of the aquatic macrophyte RA for pesticides (AMRAP) workshop (Maltby et al., 2010) for testing other macrophyte species. According to AMRAP, if an a.s. with a specific toxic mode of action (e.g. auxin inhibitors) is under evaluation for which *Lemna* may not be a representative sensitive macrophyte and/or if indications exist that terrestrial dicot species are more sensitive than terrestrial monocot species, this indicates a need for a test with a dicot aquatic macrophyte species. The PPR Panel suggests, as a pragmatic approach, to perform a test on a dicot macrophyte when terrestrial dicot species are more than a factor of 10 more sensitive than monocot terrestrial species. Following the results of the SSD analysis reported in the paper by Giddings et al. (2013), *Myriophyllum spicatum* should be a preferable dicot species to test additionally. The AMRAP document (Maltby et al., 2010) recommends the use of growth rate endpoints for macrophytes. These growth rate endpoints should preferably be based on the most sensitive ecologically relevant endpoint. In addition, a *Myriophyllum* test can also be applied to account for the exposure route via sediment. The AMRAP book (Maltby et al., 2010) recommends that if a chemical is known to partition to the sediment from the water column, and root uptake of the pesticide from sediment is likely to be an important route of exposure, an additional *Myriophyllum* test has to be performed. In case the first tier RA shows that monocot species are clearly more sensitive than dicot species and exposure via sediment is identified as an important exposure route for this compound, *Glyceria* may be a suitable test species.

7.3. Deriving regulatory acceptable concentrations

Table 23 and Table 24 summarise how to derive acute and chronic RAC values for each species.

Table 23: Endpoints available from acute aquatic toxicity tests; basic dossier data are indicated in bold (based on Commission Regulation (EU) No 283/2013 for approval of active substances).

Taxonomic group	Species/test system	Duration	Endpoint	Regulatory acceptable concentration (RAC)
Fish	***Oncorhynchus mykiss***	**96 h**	**LC_{50}**	**LC_{50}/100**
Crustaceans	***Daphnia* sp. (*D. magna* preferred)**	**48 h**	**EC_{50}**	**EC_{50}/100**
Insects/crustaceans	Additional species, e.g. *Chironomus* spp. or *Americamysis bahia* [(a)]	48 h	EC_{50}	EC_{50}/100

(a): The PPR Panel recommends to preferably use a *Chironomus* test, if not data on *A. bahia* are already available.

Table 24: Endpoints available from chronic aquatic toxicity tests; basic dossier data are indicated in bold (based on Commission Regulation (EU) No 283/2013 for approval of active substances).

Taxonomic group	Species/test system	Duration	Endpoint	Regulatory acceptable concentration (RAC)[(a)]
Fish	Early life stage test		EC_{10}(NOEC)	EC_{10}/10
Fish	Fish full life cycle test		EC_{10}(NOEC)	EC_{10}/10
Crustaceans	***Daphnia* sp. Or additional species[(b)]**	**21 d**	**EC_{10} (NOEC)**	**EC_{10}/10**
Insects	*Chironomus* spp.	20–28 d	EC_{10}(NOEC)	EC_{10}/10
Oligochaete	*Lumbriculus* spp.	28 d	EC_{10}(NOEC)	EC_{10}/10
Algae	**Green algae (e.g. *Pseudokirchneriella subcapitata*)**	**72 h[(c)]**	**E_rC_{50}**	**E_rC_{50}/10**
Algae	Diatom (e.g. *Navicula pellucilosa*) and/or blue-green algae	72 h[(c)]	E_rC_{50}	E_rC_{50}/10

Taxonomic group	Species/test system	Duration	Endpoint	Regulatory acceptable concentration (RAC)[(a)]
Macrophytes[(d)]	*Lemna* sp. or *Myriophyllum* sp. or *Glyceria maxima*	7 d–14 d	E_rC_{50}	$E_rC_{50}/10$

(a): EC_{10} is the preferred endpoint to be used. However, if the EC_{10} cannot be determined from the available data an explanation shall be provided and a NOEC may be used instead.
(b): If acute tests have been conducted on two aquatic invertebrate species, the acute endpoints shall be taken into account in order to determine the appropriate species to be tested in the chronic toxicity study.
(c): The test duration based on other Technical Guidelines, i.e. EPA OPPTS 850.4500, for algae is 96 h instead of 72 h. Endpoints from these tests are also acceptable for deriving a chronic RAC for algae.
(d): *Lemna* sp. is the default macrophyte test species for herbicidal products. In case *Lemna* and algae are apparently not sensitive to the herbicidal product (e.g. EC_{50} > 1 mg/l), or if the herbicide simulates a plant growth hormone, a rooted macrophyte is preferred (preferably *Myriophyllum*). It is adviced to test *Glyceria* in the case of a herbicide that primarily affects monocots in terrestrial plant trials.

7.4. Further testing on aquatic organisms

Additional studies may be conducted to refine the risk identified or to address additional concerns (e.g. endocrine disrupting effects). Studies should provide sufficient information and data to evaluate potential impacts on aquatic organisms under field conditions.

Additional studies undertaken can take the form of additional species testing (chapter 8), modified exposure testing (section 9.2), microcosm or mesocosm studies (section 9.3).

Where aquatic acute and/or chronic risk is indicated by the RA, expert judgement shall be used to decide the type of further assessment and additional studies required. This judgement will take into account the results of any additional data over and above those required by the present Regulation. The need to perform such studies shall be discussed with the competent authorities.

Before performing these studies, it is recommended to seek agreement of the competent authorities on the specific aims of the study to be performed and consequently on the type and conditions of the study to be performed.

7.5. Specific requirements for formulated products

7.5.1. Requirements/triggers for formulated products – acute toxicity

The requirement for formulation studies are given in Commission Regulation (EU) No 284/2013. Testing of formulated products shall be performed where:

a) the acute toxicity of the preparation cannot be predicted on the basis of the data for the a.s.; or

b) the intended use can result in direct exposure to water; or

c) the extrapolation on the basis of available data for a similar preparation is not possible.

In principle, acute or short-term exposure tests should be carried out on one species from each of the groups of tier 1 aquatic organisms (fish, aquatic invertebrates, algae and/or macrophytes) if the preparation itself may contaminate water. However, where the available information for an a.s. permits the conclusion that one of these groups is clearly more sensitive (factor of 10 difference), only a test using a species of the relevant group needs to be performed.

In addition, in the case of herbicides and plant growth regulators and other substances where there is reason to suspect effects on plants, tests should be carried out on one aquatic macrophyte species (in case several species have been tested, test on the most sensitive), if the preparation itself can contaminate water.

If the preparation contains two or more a.s., and the most sensitive taxonomic groups for the individual a.s. are not the same, testing on all tier 1 aquatic groups is required – unless a robust scientific reasoning regarding the to-be-expected mixture toxicity allows for a waiving of formulation (see section 10.3).

In order to minimise fish testing, a threshold approach should be considered for testing acute toxicity in fish (see sections 7.2.3 and 11.4).

7.5.2. Requirements/triggers for formulated products – long-term (chronic) toxicity

According to the data requirements (Commission Regulation (EU) No 284/2013), chronic studies on fish and invertebrates for formulations should only be conducted where it is not possible to extrapolate from data obtained in the corresponding studies on the a.s. (i.e. the PPP is more acutely toxic than the a.s. by a factor of 10), unless it is demonstrated that exposure will not occur. However, if the applicant demonstrates that the increased acute toxicity of the preparation is a result of co-formulants that will rapidly disappear and latency of effects is not to be expected, the RA can be based on the data for the a.s. and a chronic study with the PPP is deemed not necessary.

If chronic toxicity studies with the PPP are required, generally, studies similar to those conducted for an a.s. are required. It can be used as a higher tier option in the RA to construct an ETR using the fraction of PEC_{sw} originating from spray drift if the applicant shows that the co-formulants are not present in the other routes of exposure (i.e. run-off and drainage). However, this RA cannot be used to overrule an ETR constructed using chronic data for the a.s. and PEC_{sw} integrating all routes of exposure (i.e. the 'normal' PEC_{sw} or PEC_{max}). This may not be applicable when the formulation contains multiple a.s. (see guidance provided in section 10.3). An alternative is to conduct a specific microcosm study with the PPP to investigate long-term risks.

7.5.3. Use of formulated data in hazard and risk assessment

7.5.3.1. Comparing a.s. and formulated PPP toxicities

Where the comparison of toxicity data between the a.s. and formulated product – often limited to acute standard endpoints – reveals differences, the RA of the PPP should (in accordance with Commission Regulation (EU) No 284/2013) be based on the lower of the two EC_x (or NOEC) values (i.e. either a.s. data or formulation data are used). For a PPP with more than one a.s., a comparison of the measured formulation toxicity and calculated mixture toxicity could indicate a potential impact of co-formulants (see section 10.3).

7.5.3.2. Bridging data gaps with similar formulations

When measured data are not available for the formulation to be assessed but are available for a 'similar' one, the toxicological equivalence of formulations must be assessed. A priori, such bridging is least critical when the formulations differ slightly only in the a.s. content but not in the nature and contents of co-formulants. If the composition of the formulations differs significantly (different co-formulants and/or considerably different concentrations of identical co-formulants), bridging is not possible without detailed case-by-case evaluation and scientific justification for the eco-toxicological comparability of the formulations in question. In particular, suitable experimental bridging studies for the sensitive organism group of standard test species can demonstrate that the toxicity of two formulations is relatively similar (i.e. PPP to be assessed and PPP with a deviating composition and data available). For example, if single species tests indicate that the toxicity of the formulations does not differ for the most sensitive organism group, then data from an experimental ecosystem study conducted with the different formulation could be used. In line with the GD on the equivalence of technical materials (EC, 2012), it is assumed that a difference in toxicity is relevant to RA, if the relevant NOEC or EC_x values differ by a factor of more than three in studies with the same species conducted under comparable test conditions; however, it is only possible to make a reliable comparison of the results if the studies were carried out according to the same testing methods and under comparable exposure conditions (e.g. static, flow-trough). In addition, there can only be

bridging between such tier 1 standard groups of organisms for which it can be assumed that the effect of the test substance is based on the same mechanism of action in both groups (e.g. in the case of a herbicidal a.s., bridging is generally possible between aquatic plants, but not between plants and invertebrates). If it is not possible to prove the ecotoxicological comparability, new studies with the formulation to be assessed shall be presented.

7.6. Bioconcentration and secondary poisoning

Some compounds in the water have the tendency to accumulate in the tissue of fish or in the tissue of other organisms. This tendency of a compound is often expressed in a bioconcentration factor (BCF). The equilibrium concentration for a compound in fish can be estimated by multiplying the concentration of the compounds in the surrounding water by the fish BCF for that particular compound. At equilibrium, the BCF also equals the ratio of the uptake rate constant and depuration and elimination rate constant (Mackay, 1982).

The bioconcentration of the substance should be assessed where:

- the log P_{ow} is greater than 3 or other indications of bioconcentration (for instance monitoring data in biota or structural alerts), and
- the substance is not rapidly degraded in water (i.e. there is less than 90 % loss of the original substance over 24 hours via hydrolysis).

Note that for ionising substances, a changing environment may have effects on the lipophilicity and thus also the bioconcentration of substances. Therefore, the bioconcentration for ionising chemicals needs to be assessed on a case-by-case basis.

Also note that there are additional factors that could influence the potential of a compound to accumulate, for instance active transport, metabolism and metabolite accumulation potential, affinity to other tissues than fat, the possibility of rapid hydrolysis ($DegT_{50} \leq 12$ hours) and the molecular weight (molecules greater than 700 are difficult to be taken up by gills).

7.6.1. Bioconcentration in fish

The test should provide a steady-state BCF, uptake rate constants and depuration rate constants, incomplete excretion, metabolites formed in fish and, if available, information on organ-specific accumulation. The BCF can either be a calculated BCF based on uptake rate constants and depuration rate constants or a measured BCF in organism tissue at steady state (OECD guideline 305).

All data should be provided with confidence limits for each test substance. BCF_k (kinetic bioconcentration factor) values should be reported as growth-corrected and as lipid-normalised values (default 5 % lipid content).

Data produced for metabolism, distribution and expression of metabolites in the case of the use of a.s. in fish farming may also be relevant in addressing this point.

For strongly hydrophobic substances (log $P_{ow} > 6$) a dietary test is recommended; one of the reason for this is testing via aqueous exposure may become increasingly difficult (because the aqueous concentration cannot be maintained at a level that is considered to be sufficiently constant) and the other reason is that the exposure via the food for those strongly hydrophobic substances is the predominant route of exposure compared with the aquatic route. (OECD 305 part 2)

7.6.2. Secondary poisoning

In addition to potential effects on fish, special attention should also be paid for potential transfer of lipophilic compounds through the food chain. For organic chemicals, a log $P_{ow} \geq 3$ indicates a potential for bioaccumulation. If this condition is met, a RA for secondary poisoning should be carried

out. For the aquatic system this risk is assessed for a fish-eating bird with a body weight of 1 000 g and a fish-eating mammal with a body weight of 3 000 g.

For the stepwise approach for assessing the bioaccumulation potential, see the RA methodology according to the GD for birds and mammals (EFSA, 2009c).

7.6.3. Regulatory acceptable concentration based on biomagnification

According to the previous aquatic GD (EC, 2002a), biomagnification has to be taken into account for compounds that meet the trigger for a FFLC test, namely that the BCF (whole body based on parent compound) is > 1 000 and the elimination of radioactivity during the 14-day depuration phase in the bioconcentration study is < 95 % and the substance is stable in water or sediment ($DegT_{90}$ > 100 days). The first tier assessment for biomagnification should cover the aquatic food chain (e.g. for predatory fish) and birds and mammals feeding on aquatic organisms. The previous GD states that if these triggers are met, detailed food-chain modelling should be performed or micro-/mesocosm studies, which implicitly take into account biomagnification, should be submitted. However, the methodology for food-chain modelling as proposed in EC (2002a) is very complicated and requires a lot of input data. Furthermore, including fish in micro-/mesocosm experiments can present difficulties and needs to be carefully considered. The PPR Panel therefore recommends to further elaborate a RA methodology to better address biomagnification in the future. It is therefore proposed to consider food-chain modelling as an option for higher tier assessment.

It is therefore recommended that, as a first tier, the methodology of the technical Guidance Document (TGD) (EC, 2003) and EQS-guidance (EC, 2011b) is adopted and the RA is performed using default biomagnification factors. The TGD proposes the following factors, related to BCF and/or log P_{ow} (Table 25).

- The biomagnification factor (BMF) is defined as the relative concentration in a predatory animal compared with the concentration in its prey (BMF = $C_{predator}/C_{prey}$). The concentrations used to derive and report BMF values should, where possible, be lipid normalised.
- Value 5 is the AF used in bird and mammal RA for chronic assessments.
- The values 0.159 and 0.138 are multiplication factors based on a 1 000-g bird eating 159 g of fish per day and a 3 000-g mammal eating 415 g of fish per day.

Table 25: Default BMF values for organic substances (according to EC, 2003)

Bioconcentration factor (fish)	Biomagnification factor
< 2 000 or log P_{ow}< 3	1
2 000–5 000	2
> 5 000	10

From Table 25, it can be seen that biomagnification may be relevant for compounds with a BCF ≥ 2 000 L/kg. For these compounds, the appropriate BMF will be selected from Table 25 and the RAC_{sp} will be derived according to the following formula:

$$RAC_{SP} = \frac{NOAEL_{bird}}{5 \cdot 0.159 BCF_{fish} \cdot BMF} \text{ or } \frac{NOAEL_{mammal}}{5 \cdot 0.138 BCF_{fish} \cdot BMF}$$

where:

RAC_{sp}: regulatory acceptable concentration in water for secondary poisoning (mg/L)

NOAEL: relevant long-term no-adverse-effect level for birds or mammals (mg/kg bw per day)
BCF_{fish}: whole body bioconcentration factor in fish (L/kg)
BMF: biomagnification factor from Table 25 (kg/kg)

This RAC_{sp} should be compared with the 21-day TWA PEC in surface water:

If RAC_{sp} > 21-day TWA PEC_{sw}, no further action is required;

If RAC_{sp} < 21-day TWA PEC_{sw}, refinement is necessary.

Where the need for further refinement is triggered, a higher tier assessment should be carried out and the food-chain modelling approach of the aquatic GD (EC, 2002a) can be followed. As bioaccumulation processes often are slow and substances may be persistent, a long-term assessment is appropriate. Relevant metabolites must also be considered. For background information with regard to food-chain modelling, see Romijn et al. (1993, 1994), Traas et al. (1996), Jongbloed et al. (1996) and Luttik (2003).

Besides the two routes that are used for calculating a RAC (i.e. exposures through the food chains (fish) for birds and/or mammals, already described in the GD for birds and mammals (EFSA, 2009c)), other routes do exist. For instance, there are also fish species at the top of the same food chain (e.g. pike (*Esox lucius*)). Some guidance is provided in the former GD on Aquatic Ecotoxicology (EC, 2002a section 5.7.4). It is also expected that, for instance, for amphibians, compounds may be transferred through their food chains. Because no toxicity data are available for those species (expressed in g/kg food), it is not possible for the time being to calculate a RAC based on these food chains. It is recommended to further develop a RA scheme for biomagnification in the future.

8. Higher-tier effect assessment on the basis of laboratory toxicity tests with standard and additional species

In situations where a low risk cannot be established at lower tiers, supplementary laboratory toxicity studies may be made use of. The aim of using such data is to reduce uncertainty in the RA, mainly by addressing more realistic exposure profiles and/or better capturing of inter-species variations in sensitivity.

8.1. Additional studies from the open literature

When assessing a.s. and formulated PPPs, additional laboratory toxicity data exceeding the regulatory requirements (Commission Regulation (EU) No 283/2013 and No 284/2013) will often be available, due to the legal obligations to submit scientific peer-reviewed open literature data. This open literature data may be used in identifying additional toxicity data for relevant species, which are not captured in the standard test package. Although there are defined approaches for assessing data quality that may be adopted, there are some general principles that should be considered when assessing the reliability of additional studies from the open literature that provide data for establishing or refining RA parameters. These have been reviewed by EFSA (2011) and include statistical power; verification of measurement methods and data; control of experimental variables that could affect measurements; universality of the effects in validated test systems using relevant animal/plant strains and appropriate routes of exposure; biological plausibility of results; and uniformity among substances with similar attributes and effects (adapted from Becker et al., 2009). Detailed guidance on this is provided in EFSA (2011).

8.2. Additional species: freshwater versus marine species

It is apparent from Regulation (EC) No 1107/2009 that the estuarine and marine water bodies are parts of the environmental compartment where the level of environmental risk should be assessed, e.g. transitional, coastal and marine water are specifically mentioned in the definition of 'environment' in Article 3 of the Regulation.

The data requirements, as specified in Commission Regulation (EU) 283/2013 for the a.s. and for formulated PPPs in Commission Regulation (EU) 284/2013, cover only freshwater species, with the exception of the data requirements for *Americamysis bahia* (formerly *Mysidopsis bahia*), which is a brackish water species. The obligation to include open literature effect data (for guidance see EFSA, 2011) may, however, introduce further effect endpoints derived for brackish water or marine species.

An important question to consider is whether marine toxicity data can be used for the effect assessment in the edge-of-field surface water. From several papers it seems that the sensitivity distributions of taxonomically similar freshwater and marine species to organic PPPs do not differ significantly (Maltby et al., 2005; Brock et al., 2008; Klok et al., 2012), thus indicating that the data can be combined. Differences in sensitivity may arise as a result of the effect of test conditions (e.g. salinity) on exposure profile or when comparing taxonomically dissimilar datasets.

Risk assessment for edge-of-field surface waters should only exclude supplementary information from the open literature on relevant marine, coastal or transitional species where there is evidence of significant differences in sensitivity in these non-standard organisms that would preclude combining effects data. However, inclusion of data from such groups would first require demonstration of taxonomic and ecological relevance to edge-of-field surface water.

8.3. Geometric mean-AF approach

8.3.1. Introduction

If toxicity data from additional species are available, from valid studies not belonging to the standard test species mentioned in chapter 7, it is necessary to consider which toxicity value should be used in the RA, at least if the number of available toxicity data is not high enough to apply the SSD approach

(see section 8.4 and Table 26). According to the former guidance (EC, 2002a), 'if a considerable number of additional species was tested in valid studies, then it is possible that the AFs that are applied to the lowest toxicity value could be lowered by up to an order of magnitude'. It is not further specified how much additional data would be needed to allow for lowering the AF, and in our experience, this option is not often applied in practice. Although more species are tested and thus information on the differences in sensitivity between species is available, the RA is most often still based on the most sensitive species using the default AF. The number of species to be tested according to PPP legislation effectively sets the first tier level of protection in the effects assessment. Consequently, when more data are available and the RA is still based on the lowest value without adjusting the AF, the average level of protection may exceed the level implied by the provisions of the Regulation for the authorisation of PPPs.

8.3.2. Approaches considered by EFSA

In 2006, the EFSA PPR Panel published an opinion on the approaches to deal with additional toxicity data, taking into account that the same average level of protection should be maintained. They concluded the level of protection by applying the approach described in EC (2002a) on the basis that the former data requirements were not the same for each group of organisms and each PPP, and depended on the size of the standard deviation (see also Luttik et al., 2011). Besides the tier 1 assessment, the level of protection was not specified in the former Directive 91/414/EEC. The PPR Panel therefore developed methods that either maintain the average level of protection envisaged in Directive 91/414/EEC without specification (option 1 and 2) or that can be applied to achieve any specified level of protection (options 3 to 5) (EFSA, 2006a).

For taxa where the legislation requires only one species in the tier 1 effect assessment, and by following the approach described by EFSA (2006a), this effectively sets the level of protection in the effect assessment. When additional species are tested, the same average level of protection can be maintained by taking the geometric mean value (rather than the lowest value) and dividing this value by the current AF (i.e. option 1 of EFSA, 2006a). Where the legislation requires at least two species of the same taxonomic group, this implies a higher level of protection in the effect assessment. In this case, a different procedure is proposed when additional species are tested (i.e. option 2 of EFSA, 2006a). The minimum is then replaced by the second or third lowest toxicity value, depending on the sample size available, and divided by the current AF. Note that currently, in the new tier 1 data requirements, in most cases only one standard test species per specific taxonomic group is required (e.g. one fish species, one green alga, one crustacean, one insect, one monocotyledonous macrophyte, one dicotyledonous macrophyte), with the exception of the crustacean combination of *Daphnia* and *Americamysis bahia* for insecticides (although the PPR Panel prefers *Chironomus* above *A. bahia* as second arthropod test species). Consequently, the need to select option 2 of EFSA (2006a) is less prominent under the new data requirements.

Later research (EFSA, 2008) showed that the Geomean approach can be used not only in the case of log normal distribution of the toxicity data set, but for a wide range of distributions that are symmetric and unimodal (single peak) on a logarithmic scale, and also for asymmetric unimodal distributions where the long tail is to the left. It can be also used for asymmetric distributions with long tails to the right and for some examples of bimodal distributions, provided that the standard AF includes sufficient allowance for between-species variation in toxicity, which seems likely.

The work described above (EFSA, 2006a, 2008) is mainly based on distributions of acute toxicity data. It remains to be investigated whether the same procedure can be used for chronic toxicity data as well. NOECs and/or EC_{10} values may be over/underestimates (e.g. due to wide dose spacing and limited power to detect effects often caused by small sample size). The PPR Panel recommended, however, using the geometric mean for both acute and reproductive endpoints, but only for similar endpoints and not for a mixture of endpoints (e.g. endpoints like body weight reduction and endpoints for reproduction should not be mixed), when multiple species have been tested within a taxonomic

group (EFSA, 2006a) and this recommendation is kept for this GD. The first tier AF of 10 or 100 should be applied to this geometric mean value of available toxicity data to derive a RAC.

The PPR Panel suggests using the Geomean method only if more data are available than requested in the data requirements for the first tier. In addition, the PPR Panel suggests to base the Geomean-RAC on the taxonomic group that provides the lowest geometric mean value. If, for example, for a herbicide, EC_{50} values for three green algae, one diatom, two monocotyledonous macrophytes and one dicotyledonous macrophyte are available, the Geomean-RAC will be assessed by taking the lowest value of (1) the geometric mean $EC_{50/10}$ for the three green algae, (2) the geometric mean $EC_{50/10}$ for the two monocotyledonous macrophytes, (3) the $EC_{50/10}$ for the diatom and (4) the $EC_{50/10}$ for the dicotyledonous macrophyte.

It should be noted that 'taxonomic group' can be interpreted in different ways. For instance, crustaceans and insects represent different taxonomic groups on the phylum level but are sometimes grouped into the taxonomic group of arthropods. Other examples are merging amphibians and fish into a taxonomic group of vertebrates; merging diatoms, green algae and blue-green algae into a taxonomic group of algae; or merging algae and macrophytes into the taxonomic group of primary producers. However, any taxonomic grouping used in refined RA approaches must be fit for that purpose (i.e. scientifically justifiable). The default approach should therefore be to treat the broad taxonomic groupings at lower tiers as different groups unless scientific arguments (e.g. read-across to data-rich compounds with the same mode of action, evidence from SSDs) can be provided in support of considering them as one group.

Note, there is a possibility that the outcome of the geometric mean approach could be biased (manipulated) by introducing insensitive species. In the case of differences in sensitivity of 1 or 2 orders of magnitude (factor 10–100), an assessment of this possibility has to be made. If the most sensitive species is more than a factor of 10 (for plants and chronic tests) or 100 (for acute invertebrate and fish test) below the geometric mean of all the tested species, a weight of evidence approach should be applied. Until now, little experience exists in applying the Geomean approach in the aquatic RA. It is an important research topic to calibrate this approach with other RA approaches in the RA scheme.

8.3.3. Derivation of acute and chronic regulatory acceptable concentrations

In some cases, additional ecotoxicity data may be available, but their number is too low to apply the SSD approach (see section 8.3). For this situation, it is proposed to use the geometric mean of the available toxicity values within a taxonomic group (option 1 described above; Table 26).

Table 26: Proposal for the derivation of Geomean-RACs for aquatic organisms when a limited number of additional single species toxicity tests is available. When applying this approach scientific arguments should be given why the selected toxicity data (on which the Geomean is based) concern the same taxonomic group relevant for the RA. If more data than indicated in the second column (≥ 5 or ≥ 8) are available, the Geomean approach could still be applied, but it is recommended to preferably apply the SSD approach (see section 8.3)

Taxonomic group	Number of toxicity data for different taxa of the relevant taxonomic group	Regulatory acceptable concentration (Geomean-EC_x/AF)	Field exposure concentration (PEC)
Acute risk assessment			
Aquatic vertebrates[a]	< 5 acute LC_{50s}	Geomean $LC_{50}/100$ [d]	$PEC_{sw;max}$
Invertebrates [b]	< 8 acute EC_{50s}	Geomean $EC_{50}/100$ [d]	$PEC_{sw;max}$
Chronic risk assessment			
Aquatic vertebrates[a]	< 5 chronic EC_{10s} (or chronic NOECs)	Geomean $EC_{10}/10$ [d, e]	$PEC_{sw;max}$ or $PEC_{sw;twa}$
Invertebrates [b]	< 8 chronic EC_{10s} (or chronic NOECs)	Geomean $EC_{10}/10$ [d, e]	$PEC_{sw;max}$ or $PEC_{sw;twa}$

Taxonomic group	Number of toxicity data for different taxa of the relevant taxonomic group	Regulatory acceptable concentration (Geomean-EC_x/AF)	Field exposure concentration (PEC)
Primary producers[(c)]	< 8 EC_{50s}	Geomean EC_{50}/10 [(d, e)]	$PEC_{sw;max}$

(a): i.e. fish or amphibians.
(b): i.e. separately for crustaceans and insects in the case of insecticides, and another specific taxonomic group in the case of fungicides, unless it is demonstrated that certain taxonomic groups can be combined.
(c): i.e. separately for green algae, diatoms, blue-green algae, monocotyledonous macrophytes and dicotyledonous macrophytes in the case of herbicides or fungicides with a herbicidal mode of action, unless it is demonstrated that certain taxonomic groups can be combined. E_rC_{50s} on the basis of growth rate and the most sensitive ecologically relevant endpoint are preferred (in accordance with the relevant OECD guidelines). Yield endpoints may also be used if growth rate endpoints are not provided.
(d): Of the different taxonomic groups the lowest Geomean value is selected (e.g. the lowest value for insects or crustaceans in the case of insecticides; the lowest value for green algae, diatoms, blue-green algae, monocotyledonous macrophytes or dicotyledonous macrophytes in the case of herbicides)
(e): When applying the Geomean approach to chronic toxicity data comparable endpoints should be used within the same taxonomic group.

A benefit of the approach described in Table 26 is that all species groups are treated in the same manner. Furthermore, the method will not have to change if, in the future, more than one standard test species is required for a particular group of species.

It should be noted that the geometric mean approach can only aim at an average level of protection and cannot address possible substance-specific deviations from average patterns (particularly for a.s. with novel modes of action for which the current tier 1 procedure may not be protective enough). It is therefore necessary to consider all information (including open literature data) on substance toxicity. This can be done by comparing the Geom-RAC_{sw} with the lowest available $L(E)C_{50}$ (in the acute effect assessment scheme for animals and the chronic effect assessment scheme for plants) or $NOEC/EC_{10}$ (in the chronic effect assessment scheme for animals). If the lowest toxicity value is higher than the Geom-RAC value, it is acceptable to use the Geomean approach, otherwise a weight of evidence approach should be applied. For example, for certain neonicotinoids and insect growth regulators (and for other PPPs with very specific modes of action) the geometric mean approach should be used with caution. It is known that species sensitivity can vary widely for neonicotinoids and that within insects, ETP taxa may be several orders of magnitude more sensitive than *Daphnia* or *Chironomus* (Beketov and Liess, 2008; Roessink et al., 2013). Furthermore, if, on the basis of the toxic mode of action of the a.s. (e.g. insect growth regulators and acute tests with insects), delayed effects can be expected that are not covered by the standard duration of the acute toxicity test, the Geomean cannot be used or should be based on prolonged toxicity tests.

Finally, considerable care will need to be taken in combining data from multiple species for chronic endpoints. Aside from the uncertainty that may arise from the use of NOECs (as opposed to $L(E)C_{50}$s for acute data or other EC_x values), there is considerably more potential for pooling data that is not directly comparable biologically. This may result from endpoints that are ostensibly comparable being assessed at different life stages, after different exposure windows, or which may reflect different toxic responses. The PPR Panel recommends, when applying the Geomean approach to chronic data, that for species of the same taxonomic groups, comparable endpoints should be used.

8.4. The species sensitivity distribution (SSD) approach

8.4.1. Introduction to the species sensitivity distribution approach

Species may vary markedly in their sensitivity to PPPs. This variation in direct toxicity can be described by constructing an SSD. The SSD is a statistical distribution estimated from a sample of laboratory toxicity data and visualised as a cumulative distribution function (see Figure 6). Logistic and log normal distributions are most often used, because they require less data than distribution-free methods and are relatively easy to fit with standard statistical software (Aldenberg and Jaworska,

2000; Aldenberg et al., 2002; Van Vlaardingen et al., 2004). However, while it is typically assumed that SSDs follow a normal distribution, significant deviation from normality (whether log transformed or not) should be a trigger for trying other distributions, (e.g. Burr type III, Weibull) that may provide a better goodness of fit. Techniques such as bootstrapping have been avoided, since they do not meet the assumption of normality, but if sample size is sufficiently large then (non-)parametric bootstrapping methods may provide point estimates and confidence intervals that are fit for purpose. Please note that there are many ways of calculating 5th percentiles, but the methods presented by Aldenberg and Jaworska (2000), Aldenberg et al. (2002) and Van Vlaardingen et al. (2004) provide 5th percentiles taking into account the sample size and also allowing the calculation of the uncertainty around the calculated 5th percentile.

Note that in SSDs, all species have equal weight and thus are considered of equal importance in assessing the ecotoxicological risks.

SSDs are used to calculate the concentration at which a specified proportion of species are expected to suffer direct toxic effects. These concentrations, the hazardous concentrations, are expressed as HC_x values and represent the value that affects a specific proportion (x %) of species. For regulatory purposes, usually the HC_5 is used, the hazardous concentration to 5 % of the species tested. When compared with the first tier effect assessment on the basis of standard test species, SSDs have the advantage of making more use of the available laboratory toxicity data for a larger array of species. They describe the range of sensitivity rather than focusing on a single value, they enable estimates to be made of the proportion of the species affected at different concentrations, and they can be shown together with confidence limits showing the sampling uncertainty due to the limited number of species tested. They can be used in a deterministic RA by taking an appropriate percentile from the SSD, or in a probabilistic RA by using the whole SSD. For a detailed description of the underlying assumptions of the SSD approach, see Posthuma et al. (2002), Forbes and Calow (2002) and Brock et al. (2011). Note that the median HC_5 value is the concentration that with 50 % certainty is lower than the toxicity values (e.g. EC_{50}s or NOECs) for 95 % of the species tested, while the lower limit HC_5 provides this concentration with 95 % certainty. Note also that the RAC derived from an SSD relies on the same concept as a RAC derived from lower tiers (i.e. no unacceptable effects must occur and thus the SSD-RAC derived from the HC_5 endpoint and associated AF must ensure that the whole community is protected and not only the 95 % of species tested mentioned above).

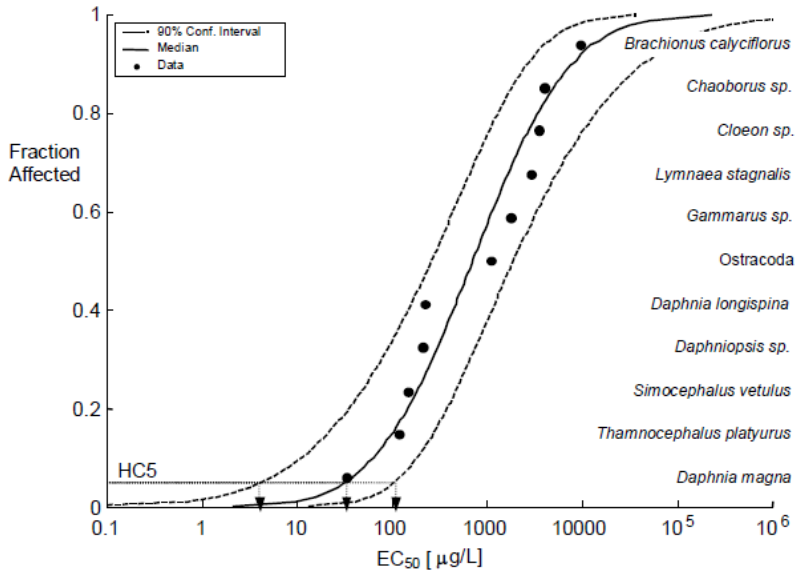


Figure 6: Graphical presentation of the species sensitivity distribution curve, its 90 % confidence interval, and as dotted arrows the derivation of the lower limit, median and upper limit hazardous concentration to 5 % of the species (HC_5). Figure from EFSA (2006c).

For construction of SSDs and the calculation of HC_5 values, several software tools are available. These programs usually also contain several statistical tools to test the assumptions of normality. It should be noted, however, that the performance of these normality tests strongly depends on the number of data. With a relatively low number of data, a distribution is often accepted as normal, whereas for large datasets deviations from normality will be more easily detected. The outcome of the tests as such should therefore not be used as a single criterion to decide whether or not the SSD approach can be applied, or to split datasets to construct specific SSDs for particular taxonomic groups (see section 8.4.3). A thorough evaluation of the individual data points and visual inspection of the fit may reveal whether or not violation of the assumptions concerning the distribution is acceptable. For example, violation of the goodness-of-fit test may be acceptable from a regulatory point of view when the fitted distribution in the tail of the SSD is relatively worst-case compared with the data points (in the sense that most of the toxicity data around the HC_5 and lower are on the right side of the fitted curve).

8.4.2. Criteria for the selection of toxicity data to construct species sensitivity distributions

The number of species data used to fit the distribution has to be adequate from a statistical point of view. Note that the 90 % confidence interval of the SSD generally will be wider, and consequently the lower limit HC_5 lower, with a lower number of toxicity data included in the SSD. The former aquatic GD (EC, 2002a) refers to the HARAP document (Campbell et al., 1999) for the aquatic effect assessment on the basis of the SSD approach. The HARAP document recommends that separate SSDs should be constructed with acute and chronic toxicity data. In addition, this document also recommends that SSDs to be based upon a minimum of either eight acute or eight chronic toxicity data for different taxa that are representative for the sensitive taxonomic group(s), at least if the SSD is not exclusively constructed with toxicity data for fish. An SSD that addresses the sensitivity of fish should be based on a minimum of 5 toxicity data points for different fish species (Campbell et al., 1999). On the basis of published SSD information for PPPs and aquatic organisms (e.g. Maltby et al., 2005, 2009; Van den Brink et al., 2006; Giddings et al., 2013), these HARAP recommendations seem to be pragmatic and appropriate from a PPP regulatory point of view. The PPR Panel recommends following these criteria.

Evaluation of scientific literature (Maltby et al., 2005; Brock et al., 2008) indicated that toxicity data from different geographical areas can be combined as long as the SSD is based on the same sensitive taxonomic group(s). It is noted, however, that toxicity studies performed in different geographic regions may be conducted under different test conditions, which may affect the exposure profile. The potential effects of test conditions on exposure should be considered whenever data are collated across different studies, irrespective of the geographical region in which the data were generated.

The endpoints measured in the toxicity tests on which the SSD is based must be the most sensitive endpoints that are toxicologically and ecologically relevant. Acute toxicity data mostly address mortality and immobility as the most frequently studied endpoints for animals, while biomass and growth rate are mostly used for primary producers. Chronic toxicity data mostly address reproduction, feeding rate and growth as the most frequently studied endpoints in animals, and again biomass and growth rate are mostly addressed by chronic toxicity data for primary producers. The use of biochemical/physiological endpoints or biomarkers in SSDs is not recommended for regulatory purposes due to difficulties in correlating results with tangible ecological effects (e.g. the protection of populations). The test duration might be a criterion to be applied for the selection of the toxicity data. Test duration, however, is taxon- and guideline-dependent and, as a consequence, a range of test durations for different organisms is often included in the same SSD.

According to Brock et al. (2011), measurement parameters, from which endpoints are calculated, should preferably be sensitive/responsive in the range of tested concentrations such that SSDs avoid the use of greater- or lower-than values. In general, it is not recommended to include unbound values (greater-than or lower-than values) in the SSD. There are situations, however, where ignoring those data would lead to a loss of valuable information. When a lower-than value is lower than the lowest toxicity endpoint, this means that the other data do not cover the whole range of sensitivities. Leaving

out this information might lead to an HC_5 that is underprotective. To demonstrate the effect of including the information in the SSD, the following procedure can be applied:

- If the greater- or lower-than value relates to a species for which a set of toxicity values is available, this value should only be included (without the < or > sign) in the calculation of the SSD for that species if it is outside the range of already available values.

- If a greater- or lower-than value relates to a species for which no other data are available, this value should only be used (without the < or > sign) in the SSD where it is outside the range of all other available toxicity values for other species or taxa.

- If an SSD is used in which unbound values are included, this should always be clearly justified.

Acute toxicity data are normally more available than chronic data due to experimental and financial constraints. In the RA for PPPs, SSDs based on chronic data are scarce. Chronic NOEC values and chronic EC_{10} effect concentrations may be included in a chronic SSD.

Whereas acute toxicity data relate to a limited number of responses and time scales, chronic toxicity data may include a wide range of responses and test durations, thereby introducing additional variability into the SSD. The test duration has to be of a chronic duration compared to the life-cycle characteristics of the species group. More specifically, a chronic toxicity test is defined as a study in which (1) the species is exposed to the PPP for at least one full life cycle or (2) the species is exposed to the PPP during one or more critical and sensitive life stages (see, for example, Holland, 1996; Brock et al., 2010a). Consequently what is considered chronic or acute is very much dependent on the species and endpoint considered. Chronic NOEC values and chronic EC_{10} effect concentrations may be included in a chronic SSD as equivalents, where there is evidence of a significant concentration–response relationship and LOEC exhibits ≥ 20 % effects (Sijm et al, 2002).

8.4.3. Selecting toxicity data on the basis of toxic mode of action of the substance

According to Campbell et al. (1999) and the former aquatic GD (EC, 2002a) the toxic mode of action of a PPP should be taken into account when constructing SSDs to derive a RAC. SSDs used in RA should always be constructed using toxicity data for the most sensitive taxonomic group (i.e. fish, invertebrates or primary producers). In the case of herbicides, vascular plants and/or algae usually consitute the most sensitive groups. For insecticides, arthropods (crustaceans and insects) usually are most sensitive. For fungicides, often a range of taxonomic groups are among the sensitive organisms.

The following information can be used to decide which taxonomic groups have to be included in an SSD for the compound under consideration. If the first tier indicates that one standard test species of the basic set is considerably more sensitive (differing by a factor > 10) than the others, in the first instance, additional toxicity data should be gathered that are representative for the sensitive taxonomic group to which this species belongs. Furthermore, data gathered by read-across on related compounds with an identical or similar toxic mode of action may give useful information on the taxonomic groups which are likely to be the most sensitive for the compound under consideration. Moreover, data in the open literature on the compound may give information on the sensitive taxonomic groups. In addition, if available, results of micro-/mesocosm tests with the compound under evaluation may shed light on the sensitive taxonomic groups, even if these tests studied the effects of relatively high concentrations not suitable to derive a threshold level of effects.

The next paragraphs give an overview of the sensitive organisms for insecticides, herbicides and fungicides (adapted after Brock et al., 2011).

8.4.3.1. Insecticide species sensitivity distributions

Evaluation of the toxicity data of 16 insecticides (including eight acetyl cholinesterase inhibitors, five pyrethroids, two organochlorine compounds and one insect growth regulator) indicates that (1)

arthropods are the preferred taxonomic group for which to construct acute SSDs, and (2) acute toxicity data for freshwater arthropods from different geographical regions and different freshwater habitats may be combined within a single SSD (Maltby et al., 2005). If necessary, toxicity data for freshwater and saltwater taxa also can be combined in an SSD, but it is important to be aware of differences in taxonomic composition and possible consequences for HC_5 values that are calculated. SSDs constructed using arthropod species recommended in test guidelines did not differ significantly from those constructed using non-recommended arthropod species (Maltby et al., 2005).

So, in the case of insecticides, arthropods (crustaceans and insects) are usually most sensitive. This implies that the SSD can focus on these taxonomic groups. Note, however, that for some novel types of insecticides (e.g. neonicotinoids) insects may be more sensitive than certain micro-crustaceans (see, for example, Beketov and Liess, 2008; Roessink et al., 2013). If, for example, the first tier toxicity value for *Chironomus* is an order of magnitude lower than that of *Daphnia* and/or *Americamysis bahia*, it is recommended to construct, in the first instance, an SSD with toxicity data for insects, or to explore which insects and crustaceans (e.g. macro-crustaceans) can be combined in a single SSD on the basis of all relevant information available. If the SSD curve constructed with toxicity data of arthropods (insects and crustaceans) fits the data well (e.g. Anderson–Darling goodness-of-fit test at $p = 0.05$), the PPR Panel proposes to preferably use the arthropod SSD in the effect assessment rather than SSD curves exclusively constructed with insects or crustaceans.

Furthermore, for certain insect growth regulators, the standard duration (48–96 hours) of the acute toxicity test may not suffice, since latency of effects may occur. It may therefore be necessary to use in the acute SSD results of prolonged acute toxicity testing on insects and/or crustaceans (e.g. by transferring the test animals to clean water after 48–96 hours exposure and continue the observations until the incipient effect level is reached).

Finally, considerable care will need to be taken in combining data from multiple species for chronic endpoints. Aside from the uncertainty that may arise from the use of NOECs (as opposed to L/EC_{50}s for acute data), there is considerably more potential for pooling data that are not directly comparable biologically. This may result from endpoints that are ostensibly comparable being assessed at different life stages, being assessed after different exposure windows, or which may reflect different toxic responses.

8.4.3.2. Herbicide species sensitivity distributions

At environmentally realistic exposure concentrations, herbicides specifically, and mainly, affect primary producers in aquatic ecosystems (i.e. algae and macrophytes). The AMRAP document (Maltby et al., 2010) and Giddings et al. (2013) provide guidance on the use of macrophyte toxicity data in the SSD approach and define areas of uncertainty which are specifically associated with the selection of species and endpoints. The uncertainty associated with species and endpoint selection to assess toxicity for algae is generally less because of the availability of standard protocols that are already used for a fairly long time.

For some types of herbicides, algae and macrophyte data may be combined in the same SSD. Van den Brink et al. (2006) and Giddings et al. (2013) showed that this is generally possible for photosynthesis inhibitors. However, herbicides that inhibit amino acid synthesis and herbicides with an auxin simulation mode of action generally seem to be more toxic to aquatic vascular plants than algae, so that it may be necessary to construct the SSD with macrophyte data (Giddings et al., 2013). Note, however, that, currently, the knowledge with respect to mode of action of several other types of herbicides is too limited to recommend the combination of algae and macrophytes, or not. If in the first tier dataset the most sensitive macrophyte is an order of magnitude lower than that of algae, a pragmatic approach is to construct in the first instance an SSD with toxicity data for macrophytes only, or to explore which aquatic vascular plant and algae can be combined in a single SSD on the basis of all relevant information available. If the SSD curve constructed with toxicity data of a wider array of primary producers (algae and macrophytes) fits the data well (e.g. Anderson–Darling

goodness-of-fit test at p = 0.05), the PPR panel proposes to preferably use the primary producer SSD in the effect assessment rather than SSD curves exclusively constructed with algae or macrophytes.

For the construction of macrophyte SSDs, the AMRAP document (Maltby et al., 2010) recommends that, in the first instance a range of morphologically and taxonomically different macrophytes should be included, unless the mode of action of the herbicide primarily affects a specific group of macrophyte species (e.g. mosses, monocotyledonous or dicotyledonous vascular plants; floating or rooted macrophytes). Ideally, SSDs should be based on toxicity values for comparable measurement endpoints generated from tests conducted under similar exposure scenarios and exposure durations, preferably using standardised protocols. The PPR Panel agrees with these recommendations and proposes to adopt them in this GD.

A more or less similar approach to that described above for aquatic macrophytes can be followed for algae. Ideally, when algae are at risk, the SSD should be constructed with a range of taxonomically different groups if the two tested algae of the tier 1 data set do not differ by more than a factor of 10 (e.g. including green algae, diatoms, blue-greens etc., and/or different genera representative for these groups).

It appears from the published literature that, for aquatic macrophytes, a wide array of measurement endpoints is used. A conservative approach would be to use the lowest endpoint per taxon, no matter what measurement parameter it is based on. Nevertheless, this wide array of available measurement endpoints may contribute to the variability in SSDs. The AMRAP document (Maltby et al., 2010) recommends the use of growth rate endpoints for macrophytes. These growth rate endpoints should preferably be based on biomass or shoot length, as they potentially provide consistency across time and species. From a statistical viewpoint, it is preferable that all endpoints used in the development of an SSD are based on common measurement parameters, since each parameter may have a different distribution. Bergtold and Dohmen (2011) present reasons why toxicity data based on specific growth rate are more informative and better suited to effect characterisation than toxicity values based on standing crop/biomass or standing crop/biomass increase (yield) for both algae and macrophytes. Moreover, according to OECD guidelines for algae (e.g. OECD guideline 201) growth rate data points are preferred. Note, however, that for mathematical reasons, an EC_{50} calculated for growth rate is usually greater than an EC_{50} calculated for biomass or biomass increase (yield) when calculated according to the OECD guidelines. The PPR Panel recommends to preferably use growth rate endpoints when both growth rate and biomass endpoints are available.

8.4.3.3. Fungicide species sensitivity distributions

Maltby et al. (2009) studied differences in sensitivity between primary producers, invertebrates and fish to fungicides with different toxic modes of action. Although for some fungicides a specific taxonomic group was most sensitive, the majority of fungicides investigated were classified as general biocides. For those fungicides that are general biocides, it is recommended to use data from all aquatic taxonomic groups to construct SSDs (Maltby et al., 2009). The HARAP document (Campbell et al., 1999) does not specify the taxonomic groups and level of taxonomic resolution when selecting toxicity data for these generic SSDs. For those fungicides that are general biocides, it is recommended as the default approach to include toxicity data from eight different taxonomic groups including at least six different orders/families in the SSD. These data include three to five toxicity data points already generated in the first tier (including fish) and five to three additional toxicity data points. For those fungicides for which a certain tier 1 taxonomic group is clearly more sensitive, it is recommended to construct, in the first instance, an SSD with toxicity data representative for this taxonomic group, if the toxicity value for this most sensitive test species is at least an order of magnitude lower than that for the other tier 1 test species. In addition, when more toxicity data are available, it is advised to always explore which taxonomic groups can be combined in a single SSD on the basis of all relevant information available. In this procedure, SSD curves are generated where a minimum of 8 data points (i.e. taxa) are available.

Initially, for fungicides with a less specific toxic mode of action towards aquatic organisms, all available aquatic toxicity data for a compound are used to generate an SSD and the fit to a log-normal distribution is assessed using the Anderson–Darling goodness-of-fit test. If the distribution does not pass the goodness-of-fit test at $p = 0.05$, separate distributions need to be constructed for vertebrates and non-vertebrates and the most sensitive distribution is used. If these distributions are not appropriately described by, for example, a log-normal model (poor fit), then the dataset is partitioned further. For example non-vertebrates may be partitioned into invertebrates and primary producers. Furthermore, invertebrates may be further partitioned into arthropods and non-arthropods, while primary producers may be further partitioned into algae and macrophytes (see Maltby et al., 2009). Note that the final SSD curve on the basis of the most sensitive taxonomic group selected should be constructed with a minimum of 8 data points (i.e. taxa) and that separate SSDs should be constructed with acute or chronic toxicity data.

It should be noted that fungi/microorganisms are not included in the standard dossier dataset as a specific taxon of interest. As a consequence, data on a potentially sensitive species group may be missing. Recent research indicates that aquatic fungi may be sensitive for ergosterol-inhibiting fungicides in particular, while for several other types of fungicide, the HC_5 based on invertebrates, primary producers and/or fish may be protective for the aquatic fungi tested (Dijksterhuis et al., 2011). The results indicate that further research into the potential effects on fungi is needed. It should be noted that the kingdom of fungi is diverse. The selection of relevant species for which standardised ecotoxicity tests may be developed should therefore be identified as a research need. Within this context it is worthwhile mentioning that several micro-/mesocosm studies with the ergosterol-inhibiting fungicide tebuconazole confirm that aquatic hyphomycetes are amongst the most sensitive endpoints (Bundschuh et al., 2011; Kosol, 2011). Micro-/mesocosm studies with other fungicides that pay attention to the responses of microorganisms at environmentally realistic exposure concentrations, however, are scarce. A recent study with the dithiocarbamate metiram revealed that aquatic fungi and bacteria are probably less sensitive than several populations of aquatic invertebrates and algae (Lin et al., 2012).

Note that if fish are included in the SSD for general biocides (non-specific fungicides), the aim is to derive a concentration that is protective at the population/community level. Since for fish a more stringent protection goal is adopted (see chapter 4), it should always be checked whether the outcome meets the regulatory lower or higher tier trigger for fish. If the SSD approach for non-specific fungicides results in a higher RAC than for fish (for example the tier 1 RAC for fish or a specific SSD for fish), the lower RAC value for fish needs to be selected as the final RAC.

8.4.4. Derivation of acute and chronic regulatory acceptable concentrations from species sensitivity distributions with invertebrates and primary producers

SSDs might be very useful in RA as they represent a cost-effective approach for the use of all available laboratory toxicity data for a larger array of species. However, unlike measures of effect generated from microcosm and mesocosm studies, toxicity values derived from acute SSDs accommodate neither recovery nor indirect, delayed or sublethal effects. With this in mind, hazardous concentrations derived from SSDs for insecticides, herbicides and fungicides were calibrated with effect class 1 and 2 data from micro-/mesocosm studies (Maltby et al., 2005; Brock et al., 2006; Van den Brink et al., 2006; Maltby et al., 2009).[32] Note, however, that these studies did not consider all PPP modes of action.

[32] The threshold concentration from micro/mesocosm studies was derived as follows: in case only effect class 1 values (see section 9.3.3.1) were available, these values were used as the threshold concentration. In case only effect class 2 were available, this value was divided by 2 to estimate the threshold concentration. When both effect class 1 and effect class 2 values were available, then the geomean of the effect class 1 and class 2 value was used as the threshold concentration. In cases where more than 1 value for effect classes was available for a compound (e.g. from different micro/mesocosm studies), then as the best approximation, the highest value was used for class 1 and the lowest value was used for class 2.

In the insecticide SSD evaluation study by Maltby et al. (2005), the majority of compounds comprised acetyl cholinesterase inhibitors (organophosphates and carbamates) and pyrethroids, while of the more novel modes of actions, only one insect growth regulator (diflubenzuron) and no neonicotinoids or biopesticides could be considered because at the time of the evaluation no appropriate micro-/mesocosm studies with these compounds were available. Recently, the calibration of the SSD approach was updated with results of micro-/mesocosm studies for insecticides by including available information on compounds with a more novel mode of action, for example, neonicotinoids or biopesticides (Van Wijngaarden et al., 2013).

For the insecticides evaluated (mainly acetyl cholinesterase inhibitors and pyrethroids), the lower limit HC_5 of acute SSDs (constructed with acute $L(E)C_{50s}$ for the most sensitive taxonomic group, e.g. freshwater arthropods) was protective for single and repeated pulse exposures in micro-/mesocosm studies, at least when the effects in these test systems were expressed in terms of nominal or measured peak concentrations (Maltby et al., 2005). This same study also states that the median HC_5 estimate based on acute toxicity data is generally – but not always – protective of single insecticide applications but that for continuous or multiple applications an AF of at least five should be applied. The recent update (Van Wijngaarden et al., 2013) also included insect growth regulators (IGRs), neonicotinoids and biopesticides in the evaluation, and showed that for 26 of the 28 insecticide cases evaluated, either the $HC_{5/3}$ or the lower limit HC_5 was protective for effects in corresponding micro-/mesocosms. The two deviating cases (one IGR and one neonicotinoid) could be explained by latency of effects (duration of the acute test apparently was not long enough to allow the expression of the effects as a result of short-term exposure).

In the SSD herbicide evaluation study (Van den Brink et al., 2006), seven of the nine compounds evaluated were photosynthesis inhibitors (six photosystem II inhibitors e.g. atrazine and metribuzin and one photosystem I inhibitor (diquat)), one compound simulated the growth hormone auxin (2,4-D) and one compound was a cell division/cell elongation inhibitor (pendimethalin). At the time of the evaluation, appropriate micro-/mesocosm studies with other types of herbicides were not yet available in the open literature. For the herbicides evaluated (mainly photosynthesis inhibitors), the lower limit of the acute HC_5 and the median value of the chronic HC_5 (based on chronic NOEC/EC_{10} values) were protective of adverse effects in aquatic microcosms and mesocosms, even under a long-term exposure regime (Van den Brink et al., 2006).

For fungicides, a larger variety of modes-of-action could be evaluated (Maltby et al., 2009), but, as discussed above, little attention was paid to effects on populations of microorganisms in fungicide-treated micro-/mesocosm tests. For fungicides, the derived $HC_{5/3}$ values, the lower limit HC_5 values and the HC_1 values were protective of adverse effects in microcosm and mesocosms studies when effects in these test systems were expressed in terms of nominal or measured peak concentrations (even under more or less long-term exposure regimes) (Maltby et al., 2009).

Maltby et al. (2009), who studied the relationship between HC_x concentrations of fungicides and corresponding threshold concentrations of effects observed in micro-/mesocosms, also reanalysed the relationships between SSDs constructed with acute toxicity data and threshold concentrations derived from microcosm and mesocosm experiments for insecticides (as published by Maltby et al., 2005) and herbicides (as published by Van den Brink et al., 2006). Maltby et al. (2005) and Van Wijngaarden et al. (2013) demonstrated that, in general (approximately with 95 % certainty), HC_5 divided by three or the lower limit HC_5 is protective of longer term exposure due to single pulse or repeated pulse applications to micro-/mesocosms. However, it should be noted that this range of AFs is not, in all cases, sufficiently protective, particularly if latency of effects due to short-term exposure may occur. For example, long-term alterations of community structure were observed in a mesocosm study after one application of 0.1 µg/l neonicotinoid thiacloprid, which corresponds to an effect concentration seven times lower than the median HC_5 LC_{50} and equals the lower limit HC_5 LC_{50} (Beketov and Liess 2008; Liess and Beketov 2011, 2012).

Furthermore, due to the nature of micro-/mesocosm studies, it should be kept in mind that the statistical power to detect effects is low as a result of the usually low number of replicates and high variability between replicates (i.e. statistically non-significant but ecologically relevant effects may occur at concentrations classified as effect threshold class 1, see minimal detectable difference (MDD) classification, see section 9.3.2.5). Furthermore effects thresholds (effect class 1 and 2 data) can vary depending on the type of data analysis (e.g. NOEC versus EC_x, taxa-based versus trait-based community-level analysis), the experimental design (e.g. spacing of tested concentrations), and the ecological characteristics (e.g. community composition) of the study. Also, even if communities in micro-/mesocosm studies would be, on the whole, representative of the range of possible communities existing in edge-of-field surface waters as considered in the meta-analyses (Brock et al., 2006), the sensitivity distribution of communities in individual micro-/mesocosm studies may not fully represent the distribution of sensitivity of communities in real edge-of-field surface waters (for more details, see section 9.3).

A default AF of 2 is therefore applied to threshold effects class 1 and an AF of 2–3 applied to effects class 2, to account for these uncertainties in extrapolating from micro-/mesocosm studies to the field situations (see chapter 9.3). Assessment factors for deriving SSD-RACs would thus need to be multiplied by this additional AF in order to maintain the calibration between these higher tiered approaches.

Table 27 presents a proposal for the derivation of a RAC for edge-of-field surface waters, based on hazardous concentrations derived from SSDs with aquatic invertebrates and plants for at least 8 different taxa belonging to the relevant sensitive taxonomic group (after Brock et al., 2011).

For primary producers, the PPR Panel recommends the calculation of the SSD-RAC both on the basis of the median HC_5 and the application of an AF of 3. This reflects the lower AF (10 instead of 100) at tier 1 for chronic data.

For invertebrates, the PPR Panel recommends the calculation of the SSD-RAC based on acute effect data on the basis of the median HC_5 and the application of an AF of 3–6.

The following aspects may be further considered in selecting the size of the AF within these ranges:

1. *The quality of the acute toxicity data used to construct the SSD.* If the toxicity data used to construct the SSD comprise several different genera/families/orders of the potential sensitive taxonomic group (see section 8.4.3 for further guidance), including EPT taxa (Ephemeroptera/Plecoptera/Trichoptera) for insecticides), a lower AF in the proposed range may be selected. However, if another valid SSD can be constructed with a more limited dataset containing the most sensitive species, and the HC_5 derived from this SSD curve is lower than that of the SSD curve using toxicity data for a wider array of taxa, a higher AF in the proposed range may be selected to be applied to the SSD from the wider set.

2. *The lower limit value of the HC_5*. If the lower limit HC_5 derived from the curve is less than one-third of the median HC_5, a higher AF in the proposed range may be warranted.

3. *The lower tier RACs on the basis of standard toxicity data (tier 1), standard and additional toxicity data (Geomean approach) and tier 3 data*. The size of the AF should ideally not result in an SSD-$RAC_{sw;ac}$ higher than the tier 3 RAC derived from effect class 1 and 2 of micro-/mesocosms studies nor lower than the tier 1 $RAC_{sw;ac}$ on the basis of standard test species and/or the Geomean-$RAC_{sw;ac}$ and/or method 3 to 5 (EFSA, 2006a) on the basis of the same toxicity data that were used to construct the SSD. Note that according to EFSA (2006a), the Geomean approach aims to achieve the same average level of protection as in the tier 1 effect assessment but can be predicted more accurately because of the availability of additional toxicity data for the relevant taxonomic groups.

4. *The position of the toxicity data in the lower tail of the SSD (around the HC_5)*. If in the lower tail the toxicity data overall are positioned on the right side of the SSD curve, the derived HC_5 estimate may be considered relatively 'conservative' for the most sensitive species. This may be a reason to adopt a lower AF from the proposed range. In contrast, if in the lower tail, the toxicity data overall are positioned on the left side of the SSD curve, this may be a reason to adopt a higher AF from the proposed range.

5. *The steepness of the SSD curve*. In the case of a relatively steep SSD curve (e.g. less than a factor of 100 between lowest and highest L(E)C_{50} value used to construct the SSD curve) a higher AF from the proposed range is recommended since exposure concentrations that exceed the $RAC_{sw;ac}$ may have ecotoxicological consequences for a larger number of taxa.

6. *Read-across information for compounds with a similar toxic mode of action*. For a PPP with a well-known mode of action sufficient higher tier information on related compounds (e.g. organophosphates) may be available that allow the evaluation of the predictive value of the median HC_5 and/or lower limit HC_5 for possible effects in micro-/mesocosms. This information may be used to select an appropriate AF within the proposed range.

7. *Considering information on chronic effects*. If the acute to chronic ratio (acute EC_{50}/chronic EC_{10}) is larger than 10, then an AF in the higher range may be warranted.

For invertebrate chronic effect assessment, the PPR Panel recommends to calculate the SSD-RAC both on the basis of the median HC_5 (and the application of an AF of 3.

Table 27: Proposal for the derivation of a RAC in edge-of-field surface waters, based on hazardous concentrations derived from SSDs with aquatic invertebrates and/or primary producers.

Type of effect/risk assessment	**Relevant predicted environmental concentrations (PEC)**	**Hazardous concentration**	**Assessment factor to derive regulatory acceptable concentration from hazardous concentration**
Acute and chronic effect/risk assessment for invertebrates and single and repeated pulse exposure	$PEC_{sw;max}$	*Latency of effects not expected.*[(a)] Median acute HC_5 (based on acute LC_{50} or EC_{50} data)[(b)]	3–6
		Latency of effects expected (e.g. insect growth regulators). Median acute HC_5 (based on acute LC_{50} or EC_{50} data from prolonged acute toxicity tests[(c)])	3–6
		or	
		Precautionary approach instead of the two options above: apply chronic SSD (see below)	
Chronic effect/risk assessment for invertebrates and long-term exposure	$PEC_{sw;max}$ or $PEC_{sw;twa}$	Median chronic HC_5 (based on chronic NOEC and/or EC_{10} data)	3

Type of effect/risk assessment	Relevant predicted environmental concentrations (PEC)	Hazardous concentration	Assessment factor to derive regulatory acceptable concentration from hazardous concentration
(concentrations during relevant time-window > 10 % of initial peak concentration)			
Effect/risk assessment for primary producers	$PEC_{sw;max}$ or, in case of macrophyte SSDs only, $PEC_{sw;twa}$	Median HC_5 (based on EC_{50} data)[(b,d)]	3

(a): This has to be demonstrated by the applicant, see further section 4.5.1. For example, by read-across for substances with similar toxic mode of action, prolonged acute toxicity tests, and information from micro-/mesocosm studies for similar compounds with a longer-term observation period after exposure.
(b): For types of PPPs evaluated by Maltby et al. (2005, 2009) and Van den Brink et al. (2006).
(c): In prolonged acute toxicity tests the observation of treatment-related responses is continued after the test organisms are transferred to clean medium.
(d): E_rC_{50}s on the basis of growth rate and the most sensitive ecologically relevant endpoint are preferred (in accordance with the relevant OECD guidelines). Yield endpoints may also be used if growth rate endpoints are not provided.

If for different taxonomic groups different and valid distributions are available, the SSD constructed with the highest number of toxicity data is used in the RA. The results of higher tier effect assessments based on SSDs have to be compared again with the results of the first tier to ensure that the RAC based on the specific SSD is protective for taxa not considered in this SSD.

8.4.5. Derivation of acute and chronic regulatory acceptable concentrations from species sensitivity distributions with fish/amphibians

The HARAP document (Campbell et al., 1999) recommends using a minimum of 5 toxicity data points to construct SSDs specific for fish. This lower number of toxicity data is chosen for, among other reasons, animal welfare considerations and because of the overall lower variability in toxicity estimates (LC_{50}; EC_{10}; NOEC) when, for example, compared with that of invertebrates. In the RA it is sometimes necessary to consider fish separately and to construct a separate SSD with fish as the most appropriate method to meet this requirement. For example, constructing a separate SSD for fish may be necessary if the risks of a PPP to populations of invertebrates and primary producers have been assessed by means of an appropriate microcosm or mesocosm experiment without fish. In regular mesocosm and microcosm studies, fish are recommended not to be included as the effects of fish might interfere with the effects of the compound on the macro-invertebrate community (Giddings et al., 2002). If potential risks to fish cannot be excluded, one appropriate method in RA is to construct a separate SSD for fish (other options are described in section 9.2).

Currently, toxicity values for aquatic stages of amphibians are not mentioned as a basic dossier requirement for the ecotoxicological effect assessment. In addition, hardly any information is available that allows a systematic comparison of the SSDs between fish and aquatic stages of amphibians. Consequently, it will depend on expert judgement whether, on the basis of the available toxicity data for fish and amphibians, a single or separate SSD has to be constructed for these taxa. When the RAC is based on a separate SSD for fish, then also a separate RAC for amphibians has to be generated to make sure they are covered in the refined RA. The separate RAC for the amphibian species may be the single species assessment approach. Whether the SSD for fish is also representative for amphibians is a topic for future research. For other options for refined vertebrate RA see section 9.2.

Acute LC_{10} and acute NOEC values may be used to construct the SSD and to calculate the HC_5 and lower limit of the confidence interval of the HC_5 ($LLHC_5$) for fish (and/or amphibians), since a higher

protection level is desired for vertebrates than for invertebrates and plants. Another option is to apply an extra AF to the HC_5 based on acute LC_{50} or EC_{50} data.

It is recommended that the following hazardous concentrations and AFs are used to derive a RAC for fish and other aquatic vertebrates (Table 28). The rationale behind the suggested AFs is an extrapolation from the AFs used for invertebrates, which has been calibrated with micro-/mesocosm experiments. However, for fish, a more stringent protection level has been adopted for the acute RA (avoiding visible mortality of individuals) and for that reason the PPR Panel proposes to apply an AF of 3 on the median HC_5 from an SSD constructed with acute NOEC/EC_{10} values for fish. In order to also derive an SSD-RAC for vertebrates based on acute LC_{50} values (since these data are usually reported in the dossiers) the PPR Panel assumes an overall difference of 3 between acute LC_{50} and acute LC_{10}/NOEC values for fish resulting in an AF of 9. For the ratio between the acute LC_{50} and chronic NOEC/L(E)C_{10}, usually a factor of 10 is assumed (see, for example, Roex et al., 2000). Taking this into account, assuming a factor of 3 for the ratio between the acute LC_{50} and acute NOEC/LC_{10} for fish seems to be appropriate. Furthermore, traditionally, an AF of 10 has been attributed to the variation in sensitivity between species (for the acute RA) and hence an AF of 9 harmonises to this assumption. Nevertheless it is acknowledged that the method proposed needs calibration.

For the chronic RA (focusing on protection of fish populations) the same reasoning and AF is proposed as for invertebrates. The PPR Panel recommends to calculate the SSD-RAC on the basis of the median HC_5 and the application of an AF (Table 28).

Table 28: Proposal for the derivation of a RAC for edge-of-field surface waters, based on hazardous concentrations derived from Species Sensitivity Distributions with fish (and other aquatic vertebrates). Information of possible latency of effect may be obtained on the basis of knowledge on the specific toxic mode of action and, read across information. To avoid unnecessary testing with aquatic vertebrates for animal welfare considerations the conduct of prolonged acute toxicity tests to demonstrate latency is not considered

Type of effect/risk assessment	**Relevant predicted environmental concentration (PEC)**	**Hazardous concentration**	**Assessment factor to derive regulatory acceptable concentration from hazardous concentration**
Acute effect/risk assessment	$PEC_{sw;max}$	*Latency of effects not expected*[(a)]. Median acute HC_5 (based on 96 h NOEC and/or acute LC_{10} data)	3
		or	
		Median acute HC_5 (based on 96h LC_{50} or EC_{50} data)	9
		or	
		If latency of effects is expected go to chronic effect assessment (see below)	
Chronic effect/risk assessment	$PEC_{sw;max}$ or $PEC_{sw;twa}$	Median chronic HC_5 (based on chronic NOEC and/or EC_{10} data)	3

(a): This has to be demonstrated by the applicant, see further section 4.5.1. For example, by read-across for substances with similar toxic mode of action, tailor-made acute toxicity tests for similar compounds with a longer-term observation period after exposure.

9. Higher-tier effect assessment by means of refined-exposure laboratory toxicity tests and experimental ecosystems

9.1. Selecting the appropriate exposure regimes when addressing time-variable exposures in higher-tier effect studies

9.1.1. Introduction

In edge-of-field surface waters, time-variable exposure concentrations of PPPs are more often the rule than the exception. The shapes of the time-variable exposures depend on factors like physical–chemical properties of the PPP, the application regime in the crop, the relative importance of different entry routes (e.g. drift, surface run-off, drainage) and properties of the receiving water bodies (e.g. water flow, water depth, pH, light penetration, biomass of plants). As an example the predicted exposure profiles for a hypothetical PPP in two stream scenarios (D1 and D5) and one ditch (D1) and pond scenario (D5) are presented in Figure 7. These predictions are based on FOCUS step 3 modelling for the use of the PPP in spring cereals (for FOCUS scenarios and model approaches see chapter 6). In this example, the predicted exposure profiles are characterised by repeated pulse exposures (predominantly owing to drift), but the peak concentrations, pulse durations and/or intervals between pulses differ between scenarios. In the exposure profiles presented in Figure 7, the $PEC_{sw,max}$ value is highest for the D1 ditch, somewhat lower in the streams (D1 and D5) and considerably lower in the D5 pond, while the pulse durations increase when going from stream, to ditch, to pond.

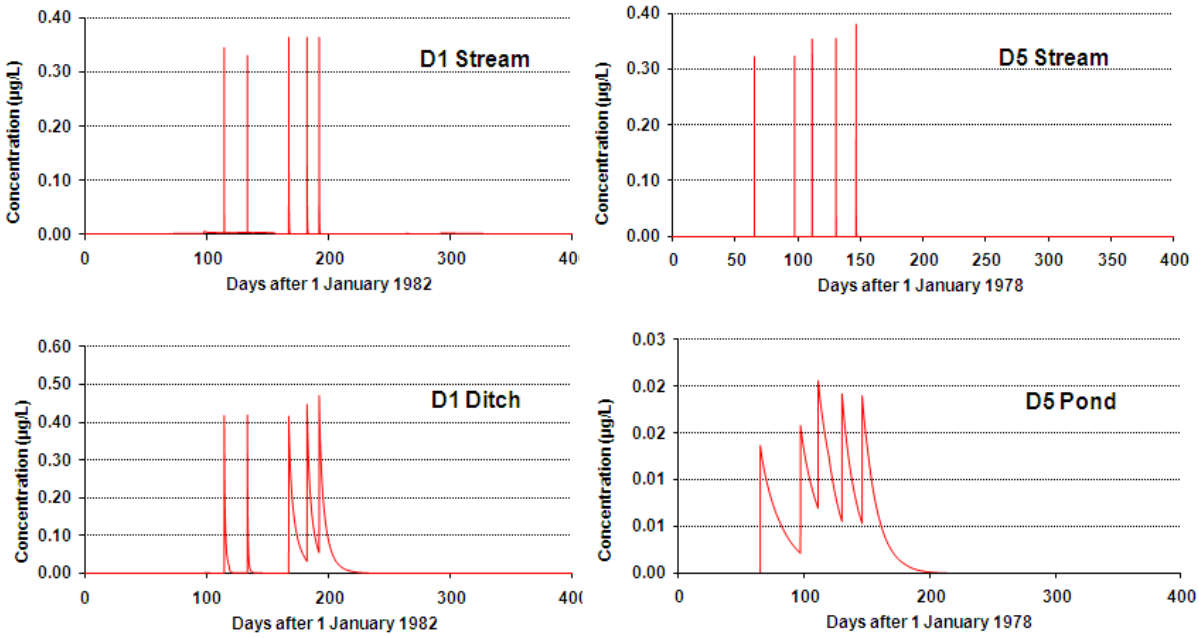


Figure 7: Predicted exposure profiles for an example PPP used in spring cereals in two stream scenarios (D1 and D5) and one ditch (D1) and pond (D5) scenario on the basis of FOCUS step 3 modelling

Higher tier effect assessments do not necessarily need to be performed by simulating constant exposures normally used in standardised lower tier tests, but may address the exposure regimes predicted for edge-of-field surface waters. For a straightforward risk and effect assessment, however, the exposure regime of the PPP in the ecotoxicological test should be realistic to worst case relative to the predicted exposure regime in the edge-of-field surface water under consideration, at least if TWA concentrations cannot be used in the RA (see section 4.5).

If the TWA concentration approach cannot be used in the chronic RA, section 9.2 describes possibilities for refined exposure single species studies and section 9.3 describes possibilities for micro-/mesocosm studies that aim to simulate realistic to worst-case time-variable exposure concentrations, in terms of height, duration, spacing and frequency of pulse exposures. The refined exposure regime tested should be guided by relevant exposure predictions for the intended agricultural uses (e.g. as deduced from FOCUS surface water scenarios or from national exposure scenarios). In the sections below, guidance is given on how to select the appropriate exposure regimes in higher tier effect studies.

9.1.2. Use of predicted exposure profiles for edge-of-field surface waters in higher tier effect assessments

Before starting a higher tier effect assessment on the basis of time-variable exposures, the predicted exposure profile for the PPP of concern in the relevant stream/ditch/pond scenario needs to be compared with the tier 1 RACs (based on standard laboratory toxicity data). In Table 29 this is done for the example PPP of which the exposure profiles are presented in Figure 7. It appears that potential acute risks are identified for the D1 and D5 stream scenarios and for the D1 ditch scenario ($PEC_{sw;max} > RAC_{sw;ac}$). For these scenarios the $PEC_{sw;max}$ is also larger than the chronic tier 1 RAC ($RAC_{sw;ch}$), however, no chronic risks are triggered ($PEC_{sw;7d\text{-}twa} < RAC_{sw;ch}$) if it is possible to use the TWA approach. For the pond scenario, no acute and chronic risks are triggered since both the acute and the chronic tier 1 RAC are higher than the $PEC_{sw;max}$.

Table 29: Comparison of PECs (peak and 7-day TWA) with the tier 1 RACs (acute: $RAC_{sw;ac}$; chronic $RAC_{sw;ch}$) for the hypothetical PPP used in spring cereals and presented in Figure 7.

Scenario	Water body	$PEC_{sw;max}$ (µg/L)	$RAC_{sw;ac}$ (µg/L)	$RAC_{sw;ch}$ (µg/L)	$PEC_{sw;7d\text{-}twa}$ (µg/L)
D1	Stream	0.36	0.20	0.17	0.03
D1	Ditch	0.47	0.20	0.17	0.14
D5	Stream	0.38	0.20	0.17	0.01
D5	Pond	0.02	0.20	0.17	0.02

Since the highest exposure concentrations (both $PEC_{sw;max}$ and $PEC_{sw;7d\text{-}twa}$) are calculated for the D1 ditch scenario, it is logical to evaluate this scenario first when selecting an appropriate exposure regime for higher tier effect studies. This can best be done by plotting the tier 1 $RAC_{sw;ac}$ (and/or tier 1 $RAC_{sw;ch}$) on the predicted D1 ditch exposure profile (Figure 8). Note that, when available, the Geom-RAC or the SSD-RAC may be used to replace the tier 1 RAC. In the example (Figure 8) the exposure profile is characterised by a repeated pulse exposure regime (five pulses) and the peaks of all pulses exceed for short periods the tier 1 $RAC_{sw;ac}$. Consequently, to evaluate the potential risks of these five pulse exposures they should be addressed realistically in higher tier effect studies, unless the number of pulses to be studied can be reduced on the basis of ecotoxicological information (for details see below).

It also needs to be checked whether the other scenarios have a higher number (with a lower $PEC_{sw;max}$ still above RAC) of peaks or if the duration of the pulses are longer in other scenarios. In the first step a worst-case exposure scenario can be constructed by selecting the maximum number of peaks, the highest $PEC_{sw;max}$ and the longest peak duration among all relevant scenarios. The number of peaks simulated in the actual test can be lowered based on the guidance below.

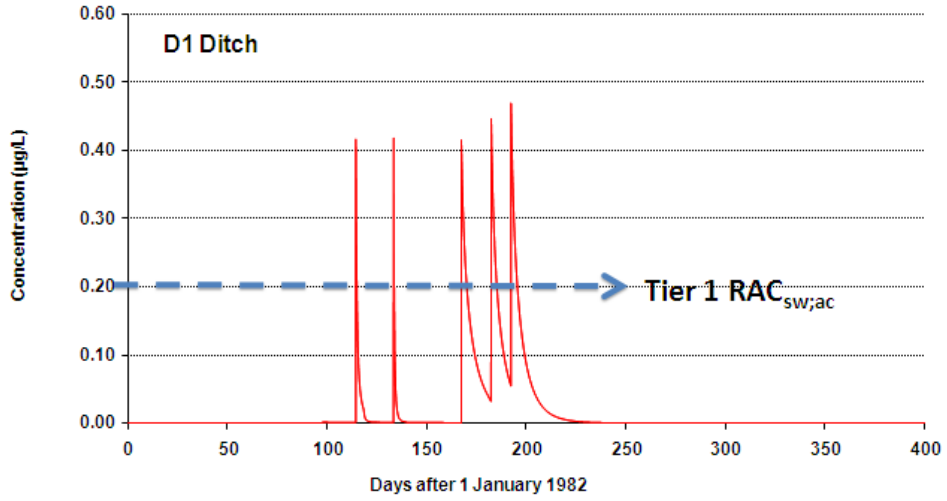


Figure 8: Exposure profile for the example PPP in the D1 ditch scenario on which the tier 1 $RAC_{sw;ac}$ is plotted.

9.1.3. Toxicological (in)dependence of different pulse exposures

For an appropriate assessment of risk from exposure profiles characterised by repeated pulsed exposures, in the first instance it is important to determine whether the pulses are toxicologically independent or not (EFSA, 2005a). Toxicological dependence of repeated pulses may occur if the life span of the individuals of the sensitive species is long enough to also experience repeated pulse exposures. For example, in the predicted exposure profile as presented in Figure 8 for the D1 ditch scenario, the three last pulses can be considered toxicologically independent from the first two pulses if the period of 32 days between the second and the third pulse is longer than the average life span of individuals, or the duration of the sensitive life stage, of the sensitive species at risk. In addition, even when individuals of the sensitive species (or sensitive life stage) have a longer life span than 32 days, the first two pulses can be considered as toxicologically independent relative to the last three pulses when:

1. the internal exposure concentrations in the individual organisms drop below critical threshold levels; and
2. complete repair of damage occurs between pulse two and three.

To demonstrate the toxicological independence of different pulse exposures, either specially designed pulsed exposure tests (section 9.2) or TK/TD models for the relevant organisms and PPP of concern may be used. If the toxicological independence of successive pulse exposures can be demonstrated for the species at risk, it may be valid to adopt a pulsed exposure regime consisting of fewer pulses in higher tier tests than present in the predicted exposure profile. If, for example, it can be demonstrated that for the species at risk the first two pulses of the example PPP in Figure 8 are toxicologically independent from the last three pulses, it may suffice to simulate the last three pulses (including the highest peak exposure) in the higher tier refined exposure effect study.

9.1.4. The minimum number of toxicologically dependent pulse exposures to address in higher tier effect studies

When addressing the ETO, it may not be necessary to incorporate all the toxicologically dependent pulsed exposures in higher tier effect studies if it can be demonstrated that fewer pulse exposures already result in the maximum effects. For example, if all pulses of the example PPP presented in Figure 8 are toxicologically dependent and the first two pulses already result in the maximum magnitude of effect (e.g. mortality), the last three pulses do not contribute to the magnitude of the response anymore although the duration of the effect probably will be prolonged by the last three pulse exposures.

An important question at stake is how to assess the minimum frequency of pulsed exposures, and the minimum duration of the time window of the pulsed exposure regime, that is likely to result in a maximum magnitude of effect. The number of PPP applications made in the higher tier effect experiment (e.g. refined exposure laboratory toxicity test; micro-/mesocosm study) has to be considered carefully in relation to the expected biological effects. However, it is considered that the number of applications in the higher tier studies should be as low as possible and is guided by:

1. the predicted exposure profile and the number and duration of toxicologically dependent pulsed exposures exceeding the lower tier $RAC_{sw;ac}$ and/or $RAC_{sw;ch}$;
2. the time-course of the responses observed in the available laboratory toxicity tests with sensitive standard and additional test species;
3. biological information of the species potentially at risk; and
4. read-across information for compounds with similar toxic modes of action.

For example, if aquatic invertebrates are at risk, the exposure period for the pulsed exposure regime need not to be longer than the overall duration of the chronic laboratory toxicity for invertebrates (21–28 days; three to four weekly pulse exposures), if it is likely that the sensitive life stages of the organisms at risk are present and the time-to-onset of maximum effect is reached in this period.

Note that in the near future, validated/calibrated TK/TD models may be used to predict the time course of effects of time-variable exposures and consequently also to identify the minimum number of toxicologically dependent pulse exposures that has to be addressed in the higher tier effect study to assure that the maximum magnitude of effects will occur.

9.1.5. Ecological (in)dependence of different pulse exposures

When the toxicological (in)dependence of successive pulse exposures is sufficiently addressed, it may be important to also demonstrate their ecological (in)dependence, particularly when ecological recovery is taken into account in the effects assessment (e.g. to address the ERO by means of micro-/mesocosm tests). Successive pulsed exposures may be considered ecologically independent if peak intervals are greater than the relevant recovery time of the sensitive populations of concern. The possible ecological independence of pulsed exposures may also be of importance in the RA if the potentially sensitive species, or specific sensitive life stages of these species, are not present in the periods that certain pulsed exposures occur (e.g. pulsed exposure in winter because of drainage).

Evaluating the ecological dependence of successive pulsed exposures will be important when microcosm and mesocosm tests are used in the RA that aim to derive the ERO-RAC (RAC on the basis of the ERO) and when the frequency of pulsed exposures (or the exposure pattern) in the edge-of-field exposure profile deviates from that tested in the micro-/mesocosm experiment (see section 9.1.4 above). In that case the total period of effects may be estimated by plotting the micro-/mesocosm derived ETO-RAC (RAC on the basis of the ETO) as well as the ERO-RAC on the field exposure profile. To illustrate this procedure for the example PPP, these RAC values are plotted on the D1 ditch exposure profile (Figure 9). The mesocosm study from which these RACs are derived is characterised by three weekly applications of the PPP (simulating the last three pulses of the D1 ditch exposure profile). The maximum magnitude of effects on sensitive invertebrate populations was observed between the second and third weekly application, while at the treatment level from which the ERO-RAC is derived full recovery of the affected populations was observed six weeks after the last application. So the ERO-RAC is based on a time window of effects of approximately seven weeks, followed by recovery. In both the upper and lower panel of Figure 9 it is assumed that the provisional ERO-RAC is 0.52 µg/L. Note that the qualifier ‘provisional’ is used since the definitive ERO-RAC can only be established when considering the total period of effects to be expected on the basis of the results of the mesocosm study characterised by three weekly pulse exposures and the total exposure profile predicted for the D1 ditch scenario. In the upper panel of Figure 9 it is assumed that the ETO-RAC derived from this mesocosm experiment is 0.43 µg/L, while it is 0.30 µg/L in the lower panel.

It appears from Figure 9 that the authorisation of the example PPP can be granted only if the SPG allows some effects followed by recovery. Evaluating the exposure period above the ETO-RAC and the time needed for recovery derived from the mesocosm test provides insight in the total effect period that might be expected. For the situation depicted in the upper panel of Figure 9 the total effect period most likely will not be longer than that derived from the mesocosm test. If an effect period of approximately seven weeks, followed by full recovery, is considered to demonstrate a low risk, then the provisional ERO-RAC can be upgraded to an official ERO-RAC. In contrast, for the situation depicted in the lower panel of Figure 9, it cannot be excluded on the basis of the provisional ERO-RAC derived from the mesocosm study (characterised by three weekly pulses) that the total period of effects will last much longer than seven weeks, since effects may already be caused by the first two pulses, while recovery most likely will not be observed earlier than six weeks after the last application. So, in the latter case, the provisional ERO-RAC cannot be used to derive an official ERO-RAC if the total period of estimated effects will be longer than acceptable from a regulatory point of view.

Further information on the interpretation of the provisional ERO-RAC will be provided in section 9.3, which deals with the conduct and interpretation of micro-/mesocosm tests. In the next chapter (section 9.2) more detailed information on the conduct and interpretation of refined exposure laboratory toxicity tests will be provided.

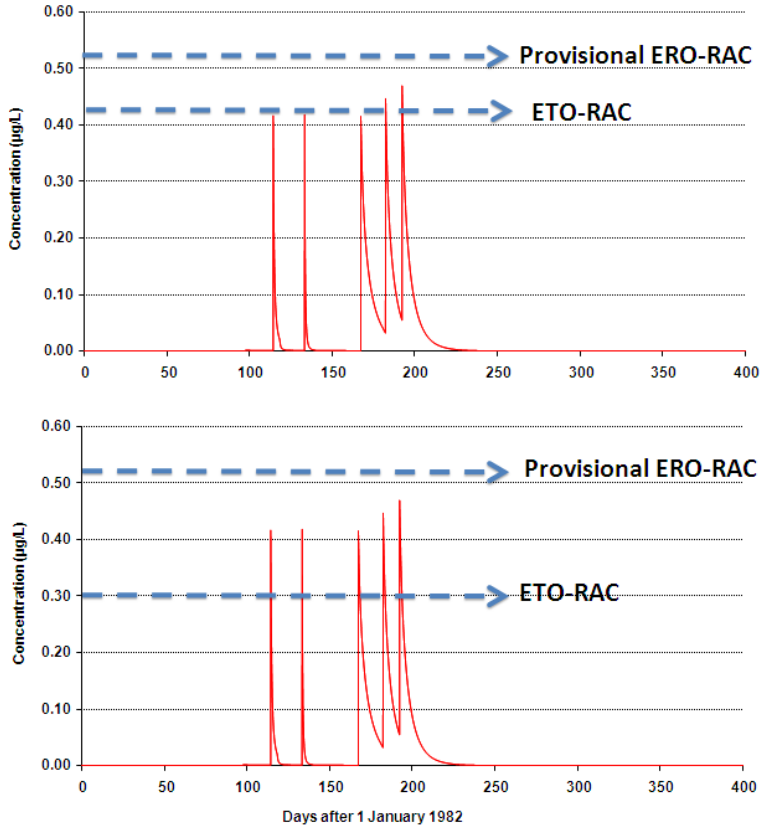


Figure 9: Exposure profile for the example PPP in the D1 ditch on which mesocosm-derived RAC values addressing recovery (ERO-RAC) or the threshold level of effects (ETO-RAC) are plotted. In the upper panel an ETO-RAC of 0.43 µg/L and in the lower panel an ETO-RAC of 0.30 µg/L is used, while the provisional ERO-RAC is the same.

9.2. Refined exposure laboratory toxicity tests

9.2.1. Introduction

RAs based on laboratory toxicity tests performed under constant exposure conditions may overestimate potential risks. In cases where predicted (modelled) field exposure profiles differ considerably from exposure regimes in standard toxicity studies it may be appropriate to design higher-tier laboratory toxicity tests that more closely resemble modelled exposure scenarios.

As described in detail in section 9.1, in designing refined exposure laboratory toxicity tests with standard and additional aquatic test species, information on the relevant field exposure predictions needs to be considered. In order to adopt a realistic worst-case exposure scenario in the toxicity test, the refined exposure regime tested should be deduced from the relevant field exposure scenarios and the relevant intended agricultural use of the PPP (e.g. FOCUS scenarios or monitoring studies). In addition, it is necessary to consider whether the tier 1 procedure triggers acute or chronic risks. In refined exposure studies supporting acute and chronic RAs the peak concentration may be used in both the PEC and the RAC estimate, if (1) the exposure profile (e.g. height and width of the pulse exposure) in the refined laboratory toxicity test (on which the RAC is based) is the worst case when compared with that of the relevant predicted (modelled) field exposure profile, and (2) the duration of the test is long enough to allow the expression of the effects.

Refined exposure laboratory toxicity tests are generally used to address the threshold level of effects and are less useful to address ecological recovery. The recovery potential of a sensitive population within a realistic community context, however, can be studied in micro-/mesocosm experiments (see section 9.3)

9.2.2. Reasons to perform refined exposure laboratory toxicity test

Performing refined exposure tests in the RA for PPPs may be done for several reasons, namely:

1. *To address the effects of time-variable exposures on relevant organisms in case it is recommended that the $PEC_{sw,max}$ is used in the chronic RA*. If the use of the TWA approach in the chronic RA is appropriate (see section 4.5), refined exposure laboratory toxicity tests need not be performed as a higher-tier option, since the $RAC_{sw,ch}$ derived from concentration–response relationships observed in the standard chronic toxicity tests, can be readily compared with the appropriate $PEC_{sw,twa}$. However, long-term refined ecotoxicological exposure studies, for example simulating repeated pulse exposures, may be a higher-tier option if the TWA approach cannot be used.

2. *In the event of doubt, to demonstrate that for the species at risk (or a representative standard test species) the $PEC_{sw,twa}$ can be used in the chronic RA*. If on the basis of the available information it is uncertain whether it is appropriate to use the $PEC_{sw,twa}$ in the chronic RA, experiments may be designed to investigate whether peak concentrations or TWA concentrations better explain the treatment-related responses observed. These experiments can be performed using different exposure scenarios that are comparable in the concentration × time factor (e.g. the same 21-day TWA concentration) but variable in peak exposure concentrations. When the response can be best expressed related to the experimental TWA concentration/area under the curve, it can be assumed that the $PEC_{sw;twa}$ approach is valid (for references of examples, see section 4.5).

3. *To address the risks of pulse exposures in addressing possible latency of effects*. When designing tests to address risks of (single or several) pulse exposures, generally the most sensitive life stage of the test organism of concern and the pulse of exposure should co-occur in the test. If it can be proven that the most sensitive life stage and the pulse exposure are not occurring at the same time, a realistic combination of exposure and appropriate life stage might be chosen. Pulsed exposure on a sensitive life stage of the test species (e.g. small

juveniles) may already yield the maximum acute and chronic effects. However, latent effects should be considered since short pulse exposures of PPPs may need a longer time to express the effect in the test organism as has been shown in the literature (Duncan et al., 2009). Hence, in the event of possible latency of effects it may be necessary that the duration of the refined exposure test simulating a short-term pulse exposure covers the full life span of the test organism, or at least the relevant sensitive aquatic life stage (see also section 8.4.5).

4. *As a higher-tier effect assessment approach for organisms that usually are not tested in micro-/mesocosms*. In principle, the effects of realistic time-variable exposure regimes on populations of freshwater organisms can be studied in aquatic micro-/mesocosm tests, if the test duration covers the full life span of the test organism as outlined above under bullet point 3 (see also section 9.3). Fish and other vertebrates such as amphibians, however, usually are not introduced into these test systems because of their undue influence on other populations (e.g. invertebrates). As a higher-tier approach, however, refined exposure tests may be used to study the individual-level effects of a realistic to worst-case time-variable exposure regime on fish (e.g. juvenile rainbow trout), amphibians and other water organisms usually not studied in micro-/mesocosms.

5. *To demonstrate the toxicological (in)dependence of repeated pulse exposures for the species at risk (or a representative standard test species)*. As discussed in section 9.1, the toxicological dependence of repeated pulsed exposures may occur if the life span of the individuals of the sensitive species is long enough to experience repeated pulsed exposures. In case of toxicological dependence a second pulse, even of a smaller height, may increase the toxicity if between pulses the internal exposure concentration has not yet dropped below the critical level and/or repair of damage has not yet occurred. To demonstrate the (in)dependence of different pulse exposures either special designed pulse exposure tests or TK/TD modelling for the relevant organism and the PPP of concern is required. Additionally, repeated pulsed exposures may occur in successive generations within a population. In such cases the overall effect on the following generations may increase.

6. *As additional information to complement results of micro-/mesocosm tests if uncertainties for a particular population remain*. Results of micro-/mesocosm experiments (see section 9.3) that did not simulate the appropriate exposure profile still may be useful for the RA if results of additional laboratory experiments (or TK/TD models) with the most sensitive species from these micro-/mesocosm experiments allow reinterpretation of the exposure–response relationships observed. In addition, if in a micro-/mesocosm experiment that simulated an appropriate time-variable exposure regime a specifically desired species is not present (or present in too low or erratic densities), additional refined-exposure experiments with this species may sufficiently complement the higher-tier effect assessment. For example, such an approach may be performed with *Lemna gibba*, since this macrophyte species usually does not grow well under the mesotrophic conditions simulated in most micro-/mesocosm experiments.

9.2.3. Refined exposure tests with standard test species

When standard test species are assessed in refined exposure laboratory toxicity tests and these tests are considered appropriate for the RA, a reduction in the AF is not justified when deriving a RAC. However, a higher toxicity value (e.g. acute EC_{50} or chronic EC_{10} or NOEC) from the refined exposure study with standard test species and the application of the standard tier 1 AF (e.g. 100 to derive the acute RAC and 10 to derive the chronic RAC) may be used in case appropriate observation times are used (prolonged for acute tests, full life cycle for chronic tests or a long-term test in the range of the life cycle).

Although refined exposure tests with standard test species that more or less resemble the design of tier 1 toxicity studies can be used for RAC derivation, the PPR Panel recommends not using refined exposure laboratory tests with populations of invertebrates (e.g. *Daphnia*) for this purpose when

recovery is also considered. These population-level laboratory experiments with invertebrates are usually performed with individuals that differ in age and developmental state. As a result a rapid onset of recovery will occur after contamination under such test scenarios. Resources for surviving individuals will increase after contamination and will trigger an unrealistic strong recovery as no competitors are present (Knillmann et al., 2012b).

9.2.4. Refined exposure tests with additional test species

The use of additional test species enables better estimation of the variability between different test species. Their inclusion in the RA is therefore highly desirable. In principle the same considerations as for standard test species are relevant. This refers to timing of exposure and latency of effect on individuals. Especially related to latency of effect, it has been reported several times in the literature that effects were only visible a long time after exposure (Duncan et al., 2009). For example, caddis flies exposed for 1 hour to the pyrethroid fenvalerate showed a NOEC of 10 µg/L after 15 days following exposure, but revealed a mortality NOEC of 0.01 µg/L eight months after exposure (Liess, 2002). A similar observation was made for the Ephemeroptera *Cloen* sp. exposed for 1 hour to fenvalerate. Four days after contamination the NOEC was observed at 1 µg/L; 29 days after contamination the NOEC related to survival was 0.001 µg/L (Beketov and Liess, 2005). When compared with the long-term effects of fenvalerate on *Daphnia magna,* these results show the high variability between test species. Additional test species should therefore be selected on the basis of a literature research identifying which groups of species are often showing latency of effect. To show latency, ideally refined exposure tests should cover the whole life cycle of the test species. These tests may be shorter if they include the most sensitive life stage of the test species (to be demonstrated by the applicant).

9.2.5. Derivation of RAC and calibration of refined exposure laboratory toxicity tests

For the derivation of an acute RAC by means of refined acute toxicity tests (which usually should have a longer duration than the standard protocol tests) with relevant standard test species, it is proposed to apply an AF of 100 (for invertebrates and fish) to the LC_{50} or EC_{50} (expressed in terms of peak concentration) under the conditions that:

- The pulse exposure in the refined acute laboratory toxicity test is realistic to worst case when compared with the relevant predicted (modelled) field exposure profile.
- The repeated pulse exposures predicted for the field are considered to be toxicologically independent (section 9.1.3). If not, the repeated pulses should be addressed in the refined and prolonged acute toxicity test.
- The duration of the acute test is long enough to allow the observation of delayed effects.
- The refined acute RAC is compared with the $PEC_{sw;max}$.

Long-term refined exposure tests with standard test species (e.g. simulating repeated pulse exposures) may be a higher-tier option if the TWA approach cannot be used. For the derivation of a chronic RAC by means of refined chronic toxicity tests it is proposed to apply an AF of 10 to the EC_{10} (or NOEC) for invertebrates and fish, or to the EC_{50} for plants, expressed in terms of nominal (if measured peak exposures do not deviate more than 20 % from nominal) or measured peak concentration in the test systems under the conditions that:

- The (repeated pulsed) exposure regime in the refined laboratory toxicity test is realistic to worst case when compared with the relevant predicted (modelled) field exposure profile.
- The duration of the test is long enough to allow the observation of delayed effects.
- The refined chronic RAC is compared with the $PEC_{sw;max}$.

A summary of the $RAC_{sw;ac}$ and $RAC_{sw;ch}$ derivation on the basis of refined exposure laboratory tests with standard test species, and their use in the RA, is presented in Table 30.

Table 30: Derivation of RACs in edge-of-field surface waters, based on refined exposure laboratory toxicity tests with standard test species

Type of effect/risk assessment	Relevant PEC	Endpoint of refined exposure toxicity test with standard test species expressed in terms of peak exposure concentration in test system	RAC
Acute effect/risk assessment	$PEC_{sw;max}$	$L(E)C_{50}$ (animal tests)	$L(E)C_{50}/100$
Chronic effect/risk assessment	$PEC_{sw;max}$	EC_{50} (plant tests)	$EC_{50}/10$
		EC_{10}/NOEC (animal tests)	$EC_{10}/10$

If a refined (prolonged) exposure test with rainbow trout (*O. mykiss*) is performed to derive a higher-tier RAC_{sw}, this RAC_{sw} may also be used to assess the effects on the larval stages of amphibians, since rainbow trout is demonstrated to be a good surrogate species for them (see section 7.2.4 and Appendix C).

In case additional test species are used in refined exposure tests, selection of test species needs to consider typical species assemblages found in undisturbed water bodies and species sensitive to the selected compound. When refined exposure studies with several additional test species of the relevant taxonomic group are available the derived toxicity values might be used as described in sections 8.3 (Geomean method) and 8.4 (SSD method), at least when conditions as described above for the derivation of refined RACs are met.

Note that the PPR Panel proposal predominantly addresses the uncertainty of the ecotoxicological endpoint. It is assumed that the predicted field exposure profile is sufficiently realistic to worst case. With regard to the calibration of assessment it is considered that long-term mesocosm experiments containing a realistic community are suitable to validate and calibrate assessments based on refined exposure laboratory tests. For this, the most sensitive population- or community-level endpoint should be used (see section 9.3).

9.3. Model ecosystem experiments

9.3.1. Introduction

Aquatic model ecosystems—usually referred to as microcosms and mesocosms—are bounded systems that are constructed artificially with samples from, or portions of, natural aquatic ecosystems, or that consist of enclosed parts of natural surface waters. They usually are characterised by a reduction in size and complexity when compared with their natural counterparts but they include an assemblage of organisms representing several trophic levels. Indoor experimental ecosystems are often referred to as microcosms and outdoor experimental ecosystems as mesocosms, but their difference mainly concerns their size. Crossland et al. (1993) defined outdoor microcosms as experimental tanks/ponds less than 15 m^3 water volume or experimental streams less than 15 m length, and mesocosms as test systems greater than 15 m^3 water volume or 15 m length. The most frequently used freshwater model ecosystems in European PPP RA are those that mimic shallow, static freshwater habitats (see Brock and Budde, 1994; Caquet et al., 2000), but in recent years more ecotoxicological experiments with PPPs in artificial streams have become available (e.g. Heckmann and Friberg, 2005; Mohr et al., 2007, 2008; Liess and Beketov, 2011).

Micro- and mesocosm studies performed for PPP authorisation aim to simulate realistic natural conditions and environmentally realistic PPP exposure regimes. These studies normally follow

experimental designs to demonstrate causality between treatment and effects, and can also identify concentration–effect relationships at the population and community level (including structural and functional endpoints). Brown et al. (2009) provide an overview of available and potential higher-tier testing approaches for the effect assessment of PPPs, including field and semi-field tests.

The advantage of micro- and mesocosm studies over the other types of experimental higher-tier studies (e.g. additional laboratory toxicity tests to construct SSDs; refined exposure studies) is their ability to integrate more or less realistic exposure regimes with the long-term assessment of endpoints at higher levels of biological integration (population- and community-level effects), and to study intra- and inter-species interactions and indirect effects in a more or less realistic community. In addition, a higher number of species and ecological groups are exposed for which dose–response relationships may be obtained. Since micro-/mesocosm tests can be performed for a relatively long time, and observations can go on long after the exposure has declined below the threshold level of effects, these test systems may be used to assess latency of effects and population and community recovery. The advantage of micro- and mesocosm studies over field monitoring studies is that owing to increased control over confounding factors, causality between PPP exposure and effects is easier to demonstrate. In addition, this kind of study allows replications and real controls, which would be not possible in a field study.

It is important to note that communities and environmental condition in micro-/mesocosms represent only one of the many possible conditions of edge-of-field surface waters. Edge-of-field surface water bodies potentially at risk vary in community structure (including species composition and life cycle traits) and abiotic conditions. This should be accounted for in the effect assessment, e.g. by applying an appropriate AF for spatio-temporal extrapolation of the concentration–response relationships observed in micro-/mesocosms.

9.3.2. Designing micro-/mesocosm experiments

Useful guidance on experimental design and endpoint selection to conduct a proper micro-/mesocosm experiment is provided by OECD (2006) and by workshop documents produced by the Society for Environmental Toxicology and Chemistry (SETAC) Europe (e.g. Giddings et al., 2002; Brock et al., 2010a; Maltby et al., 2010). The major items for an appropriate design of an aquatic micro-/mesocosm test within the context of this GD (focus on individual PPPs in edge-of-field surface waters) concern:

1. the establishment of an aquatic community in the test systems that is representative for edge-of-field surface waters and can be used in the effect assessment;

2. the appropriate exposure regime of the PPP under evaluation to simulate in the micro-/mesocosm test system;

3. the number of treatments, choice of the doses and replicate test systems per treatment to derive a concentration–response relationship;

4. the selection of measurement endpoints (type, number and frequency) for the higher-tier effect assessment; and

5. the methods for statistical and ecological evaluation of concentration–response relationships. More detailed guidance on these items is provided below.

9.3.2.1. Establishment of a representative aquatic community in the test systems

Communities used in micro- and mesocosm studies should be representative for edge-of-field surface waters. A representative freshwater community for edge-of-field surface waters contains the important taxonomical groups (not necessarily the same species), trophic groups (e.g. primary producers, detritivores, herbivores, carnivores) and ecological traits (particularly life-cycle characteristics related to vulnerability as generation time and dispersal ability, see Liess et al., 2008) typical for communities

in ponds, ditches and/or streams. Consulting ecological scenarios for edge-of-field surface waters of European landscapes may be important and some preliminary guidance on this is provided in chapter 15 of the ELINK document (Wogram, 2010; Brock et al., 2010c; Alonso Prados and Novillo-Villajos, 2010; Biggs and Brown, 2010) and in Gergs et al. (2011) and in section 5.6 on vulnerable species.

Micro-/mesocosm studies can be performed in artificial constructions (mimicking ponds, ditches or streams) or by enclosing parts of existing aquatic ecosystems (enclosures). Already established, uncontaminated, aquatic ecosystems that resemble the required species composition of the micro-/mesocosm test can be used as a source of water, sediment and organisms to seed the artificial test systems. This will ensure that into the test systems a more or less similar and representative community (e.g. characterised by zoo- and phytoplankton, pelagic and benthic macro-invertebrates, and, if necessary, macrophytes) will be introduced. It may be necessary and appropriate to additionally add certain organisms (e.g. potentially sensitive or vulnerable macro-invertebrates or macrophytes that are not present in the 'established source ecosystem') from other sources if lower tiers or other information indicate potential risks to specific organisms. It may also be appropriate to use artificial sediment and water in the tests systems if, for example, the focus of the micro-/mesocosm study is on a specific group of organisms such as aquatic macrophytes (e.g. a herbicide study in an outdoor test system with potted plants). Using micro–/mesocosms constructed with artificial sediment and water and that aim to study the effects of PPPs on invertebrates, however, may require a longer acclimatisation period to develop a realistic pelagic or benthic community.

Artificially constructed model ecosystems require a pre-treatment period of at least several weeks (plankton-dominated systems) to several months or longer (structurally more complex systems dominated by long-living macro-invertebrates and macrophytes) in order to allow the establishment of a community that has recovered from the 'construction stress', adapted to the conditions in the test system and characterised by realistic food web interactions. This will depend on the generation time of the species considered most relevant (Barnthouse, 2004) and may require a time span of few generation times (Giddings et al., 2002; Liess et al., 2006). However, if during the set-up of the test system, care is taken to introduce the organisms at natural densities, the acclimatisation period may be shorter. Currently, there are no micro-/mesocosm studies comparing directly the sensitivity to PPPs of artificially constructed systems with contrasting periods of establishment. Note that during the 'acclimatisation phase' of micro-/mesocosms other organisms (e.g. aquatic insects), originally not introduced, may colonise the test systems. This should be considered a normal ecological phenomenon, provided that the representativeness to edge of field surface water is maintained.

To adequately study potential population and community-level effects for regulatory purposes, it is for the experimental design of micro-/mesocosm tests important that enough representatives of potential sensitive (and vulnerable) populations are present. A relevant question at stake is: what should be the minimum number of potential sensitive/vulnerable populations in micro-/mesocosms from which concentration–response relationships can be derived to make the study useful for higher-tier RAC derivation? Note that for RAC derivation that addresses the ETO it is the sensitive populations that need most attention, while for RAC derivation addressing the ERO most relevant are species that are both sensitive and have a slow recovery potential (e.g. due to a long life cycle and/or poorly developed dispersal abilities). In theory, the required number and identity of potentially sensitive/vulnerable populations that needs to be present in a micro-/mesocosm test system will depend on prior knowledge on the level of sensitivity/vulnerability of these populations to the PPP of concern. If, for example, laboratory toxicity tests indicate that certain species are among the most sensitive tested (e.g. situated in the tail of SSDs), and individuals of these species are sufficiently present in the micro-/mesocosm test, the threshold level for toxic effects derived from these tests may be surrounded by less uncertainty than when this knowledge is not available. However, species that are identified in laboratory SSDs as most sensitive need not be native and probably will not occur in micro-/mesocosms constructed with components of natural ecosystems. In practice, we often do not know which native species constitute the most sensitive ones. In contrast it is often known what constitutes the potentially sensitive taxonomic group (e.g. on the basis of lower tiers and read-across for compounds with a similar toxic mode of action). In these cases it seems fair that, besides

representatives of different trophic levels, at least 8 different populations of the sensitive taxonomic group need to be present in the micro-/mesocosm test systems and for which a concentration–response relationship can be derived. Note that also when applying the SSD approach, laboratory toxicity data for a minimum number of 8 different taxa of the sensitive taxonomic group are required in the effect assessment of PPPs for invertebrates and/or primary producers. In line with the data requirements for the SSD approach the sensitive taxonomic group may comprise crustaceans and/or insects for insecticides, algae and/or macrophytes for herbicides and a wider array of non-vertebrate taxa for fungicides.

Particularly when adopting the ecological recovery concept to derive a RAC it should be carefully evaluated whether the potentially vulnerable taxa in edge-of-field surface waters (e.g. sensitive univoltine and semivoltine invertebrates with a low dispersal ability or macrophytes with a relatively slow growth rate) are sufficiently represented in the test system (Appendix D). If not, it may be necessary to apply a higher AF to extrapolate the study specific NOEAEC (no observed ecologically adverse effect concentration), to use modelling tools to extrapolate the observed rate of recovery to that of more vulnerable field populations, or to derive the RAC on the basis of the ecological threshold concept. Note that species sensitivity distributions constructed with acute EC_{50s} for aquatic arthropods and insecticides suggest that the sensitivity of aquatic insects is not correlated with voltinism of the species involved (see, for example, Brock et al., 2010c). Although short-cyclic insects may have similar sensitivity distributions as insects with a more complex life cycle, it is frequently reported that sensitive multi-/bivoltine insects recover faster from insecticide stress than sensitive uni-/semivoltine insects (e.g. Van den Brink et al., 1996; Brock et al., 2009; Liess and Von der Ohe, 2005). Similarly, sensitive short-cyclic algae may recover faster from herbicide stress than sensitive macrophytes with a slower growth rate (e.g. Coors et al., 2006). If populations of invertebrates and macrophytes characterised by a low recovery potential are identified to be at risk, efforts should be undertaken to introduce representatives of these populations when constructing the micro-/mesocosm test systems. In the event that populations of these species are affected, the chance that they will recover within 8 weeks is small. Hence, in most cases where uni/semivoltine invertebrates and/or slow-growing macrophytes are identified as sensitive groups the recovery option will not bring us further and the proposed scheme (Figure 10) will direct to the threshold option in the effect assessment. When adopting the threshold concept to derive a RAC it is possible to base it on negligible effect concentrations for sensitive taxa with short generation times as sensitivity and generation time seem to be not linked. Hence to apply the threshold concept it needs not to be a problem when sensitive univoltine and semivoltine invertebrates with a low dispersal ability or macrophytes with a relatively slow growth rate are not sufficiently represented in the test systems. Instead, the availability of data on negligible effect concentrations for species sensitive to PPPs (high toxicological sensitivity) may suffice to derive an ETO-RAC (RAC that addresses the ETO).

When invertebrate and/or plant communities are the principal endpoint of the study, it is recommended that free-living fish are not included (Giddings et al., 2002). In smaller micro-/mesocosms fish usually cannot be introduced at natural biomass levels appropriate to the abundance of their prey, and therefore fish can have an undue influence on other populations inhabiting these confined test systems. However, separate micro-/mesocosms may be used to study the individual-level effects of a realistic exposure regime on fish (e.g. juvenile rainbow trout, sticklebacks (see further section 9.2 on higher tier refined exposure tests)

9.3.2.2. Selection and characterisation of the exposure regime

It is important to evaluate the possible exposure regimes in relevant edge-of-field surface waters that may result from normal agricultural use of the PPP of concern (e.g. by FOCUS surface water scenarios and models, multi-year application; see chapter 6), and to identify the relevant exposure regimes that should be addressed in the acute or chronic effect assessment (see section 9.1). If the micro-/mesocosm test is triggered by the tier 1 acute RA and the expected and relevant field exposure regime is characterised by a single high pulse (e.g. owing to drift application), or by repeated pulses that are toxicologically independent (for criteria, see section 9.1.3) a single application experimental design is

an appropriate exposure regime to study in the micro-/mesocosm experiment. The pulse duration should either be equal to or larger in the micro-/mesocosm experiment than that predicted for the field, giving a realistic to worst-case estimation, respectively. In these cases, the $RAC_{sw;ac}$ derived from the concentration–response relationships in the micro-/mesocosm tests can be expressed in terms of the peak concentration of the PPP in the test systems and this RAC estimate can be compared with the $PEC_{sw;max}$.

The nominal concentration can be used in the acute effect assessment if, shortly after application, the measured exposure concentrations in the integrated water column of the test system do not deviate more than 20 % from nominal. Note that during the first hours post application, a heterogeneous distribution of the test compound in the water column is common which may hamper the proper measurement of peak concentrations. For fast-dissipating compounds the proper measurement of the actual peak concentration in the test system may be difficult if not performed shortly after application. An alternative option to estimate the peak concentration in the test systems may be to measure the concentration in the application solutions, as well as the amounts of application solution applied to each test system. In repeated application studies the peak concentration may occur immediately after the last application if the compound does not dissipate completely from the water column between applications.

If the expected exposure regime in the field triggers concerns regarding repeated pulsed exposure that are probably considered to be toxicologically dependent (see section 9.1.3), a repeated exposure regime should be adopted in the micro-/mesocosm experiment to determine a $RAC_{sw,ac}$ and/or $RAC_{sw,ch}$ (dependent on the tier 1 RA that triggered the micro-/mesocosm study). Guidance on the minimum number of toxicological dependent pulse exposures to address in the micro-/mesocosm experiment is provided in section 9.1.4.

In a micro-/mesocosm experiment triggered by the tier 1 chronic core data, a worst-case approach is to maintain a more or less constant PPP concentration for at least the duration of the chronic toxicity test that triggered the micro-/mesocosm test, unless the TWA exposure can be used to express the treatment-related effects (see section 4.5). Alternatively, the long-term exposure regime simulated in the test systems should be realistic to worst case relative to the predicted exposure profile.

To appropriately characterise the exposure regime in the micro-/mesocosm experiment it generally is not sufficient to report nominal concentrations only. In addition to the dissipation DT_{50} in the experiments and the timing of application the following information has to be provided. The analytical recovery of the test substance in the relevant matrix (water within the context of the current Aquatic Guidance), exposure concentrations of the dosing solutions used to treat the test systems, and concentrations measured in water (and sediment) samples at several time points after each application, should be reported. This allows the expression of the effect estimates (e.g. NOECs; EC_x values) in terms of the ecotoxicologically relevant concentration (ERC; see section 5.2).

9.3.2.3. Number of treatments, choice of the doses and replicate test systems per treatment

Because of the ecological complexity of micro-/mesocosm experiments it should be realised that practicality (in terms of manpower and costs) will limit the number of test systems that can be constructed, treated, sampled and analysed. According to OECD (2006) the number of treatments and choice of doses, as well as the number of replicate micro-/mesocosm test systems per treatment depend on the objectives of the study. These objectives may relate to the required effect endpoint (e.g. EC_{50}, EC_x, NOEC values for relevant endpoints), the level of precision required for the effect estimates (desired power of the test) and the size of the effect which is considered of ecological significance. For regulatory purposes of PPPs the delegates of the CLASSIC workshop (Giddings et al., 2002) and OECD (2006) recommended an exposure–response experimental design with preferably five or more concentrations, and at least two, but preferably more, replicates per concentration. An exposure–response experimental design characterised by a larger number of treatments but a lower number of replicates per treatment, allows wider use of the data under different regulatory conditions

(e.g. different exposure regimes due to differences in application and mitigation methods, receiving water bodies etc.) than an ANOVA design with a limited number of treatments but a high number of replicates per treatment. When adopting an exposure–response experimental design it is common practice to select more replicates for the controls (often double the amount) than for treatments, not only to increase the statistical power but also as a back-up to avoid (statistical) evaluation problems in case control test systems are lost in the course of the experiment (due to human errors and/or demonic effects). Preferably, the lowest test concentration should not result in treatment-related responses, while the highest concentration tested should result in pronounced effects on several measurement endpoints. This allows the derivation of threshold concentrations for toxic effects, as well as putting in perspective the possibly more subtle treatment-related responses caused by intermediate concentration levels. This implies that the selected exposure concentrations should always be guided by lower-tier effect information (e.g. single species toxicity tests) and the expected field exposure regime of the substance under evaluation. For this purpose TK/TD and/or population models might be used; some models available nowadays are already in a state to provide this information. Validation and testing of the models itself is not as crucial for designing a mesocosm experiment as for the direct application in RA. It is expected that using these model approaches will come up with more precise studies than expert judgement.

9.3.2.4. Measurement endpoints

The appropriate measurement endpoints to study, and the frequency of sampling, in a micro-/mesocosm experiment will depend on the SPGs defined for the water organisms potentially at risk in edge-of-field surface waters (see chapter 5). According to the SPGs defined, the key drivers and their ecological entity to be protected concern populations in the case of aquatic algae, vascular plants and invertebrates, individual populations in the case of aquatic vertebrates and population functional groups in the case of aquatic microbes. This implies that for most organisms at risk that are studied in micro-/mesocosm tests the selected measurement endpoints should relate to relevant population-level endpoints, more specifically the attributes survival/growth and abundance/biomass (see section 5.5). By protecting sensitive/vulnerable populations of primary producers and invertebrates, community-level effects (including processes) and biodiversity may be sufficiently ensured, particularly when addressing the ETO.

The duration of the study and frequency of sampling should be adapted to the treatment regime (e.g. more frequent sampling immediately after PPP application), the duration of the life cycle of the organisms of concern (e.g. initially more frequent sampling for organisms with a short life cycle), and the objective of the study (e.g. if the study aims to demonstrate recovery or not and the pre-defined period of acceptable effects). In all cases pre-treatment samples should be taken (at least one or two pre-treatment samplings) to demonstrate the suitability of the test systems in terms of similarity of relevant parameters between test units. For detailed information on test duration and sampling, refer to OECD (2006) and the CLASSIC document (Giddings et al., 2002).

Microcosm and mesocosm experiments are test systems that allow the study of treatment-related effects of PPPs at the population and community level. Population responses in the micro-/mesocosm are usually studied by means of measurement endpoints that provide information on dynamics in population abundance, biomass and/or growth. Measurement endpoints to study community-level responses usually comprise community metabolism endpoints indicative for ecosystem processes such as the dynamics of dissolved oxygen in water and rates of decomposition of particulate organic matter (e.g. in litter bags), but also summary parameters of population-level endpoints, such as diversity indices and scores based on multivariate techniques and trait-based groupings of organisms (for more details see section 9.3.2.5).

The number of taxa/populations occurring in micro-/mesocosms, and consequently the potential measurement endpoints, may be high. Studying all potential measurement endpoints is very expensive. For reasons of cost-effectiveness usually a limited number of measurement endpoints is selected. Available lower-tier studies for the PPP under evaluation (e.g. standard and additional laboratory

toxicity tests) and/or results of model ecosystem experiments with related compounds (characterised by a similar toxic mode of action) may provide insight which structural and functional parameters should be studied intensively. For example, if the PPP under investigation is a selective herbicide and the laboratory toxicity tests indicate that algae and the macrophytes *Lemna* and/or *Myriophyllum* are at least an order of magnitude more sensitive than the invertebrates *Daphnia* and/or *Chironomus*, the primary focus of the selected measurement endpoints should be on populations of phytoplankton, periphyton and macrophytes (structural endpoints for primary producers) and possibly also on parameters indicative for the functioning of primary producers, such as dissolved oxygen and pH. This does not mean that no attention should be paid to responses of invertebrates (e.g. to demonstrate indirect effects) but that it may be enough to select a limited number of measurement endpoints to monitor the treatment-related effects on taxonomic groups that probably will not be impacted because of direct toxic effects. If the PPP of concern is an insecticide for which standard toxicity tests and model ecosystem experiments with related compounds indicate that crustaceans and insects in particular are sensitive, the focus of the study should be on populations of zooplankton and macro-invertebrates (possibly including emergent insects and the effects of shredder populations on the breakdown of particulate organic matter) while a limited number of (summary) measurement endpoints for primary producers may be sufficient. However, if recovery of indirectly affected organisms is of concern it may be necessary to study endpoints related to this response too in more detail.

In contrast, if the difference in toxicity between the standard test organisms is small, as might be the case for fungicides with a biocidal mode of action, the selected measurement endpoints should include a variety of taxonomic groups such as populations of primary producers (e.g. algae and vascular plants) and invertebrates (e.g. zooplankton and macro-invertebrates, including non-arthropods) and when deemed relevant (e.g. for triazole fungicides; see section 5.5), also microorganisms. Note, however, that currently validated and internationally recommended (e.g. OECD) methods to study population-level effects of PPPs on microorganisms are scarce. The available toxicity tests with microorganisms usually focus on biomass or processes (Van Beelen, 2003) or concern community-level tests based on molecular methods such as sequencing techniques (Diepens et al., 2013). A way forward for edge-of-field RA of fungicides may be to study the consequences of previously exposed plant litter on the feeding behaviour and survival of shredders (e.g. Bundschuh et al., 2011). Note that leaf-shredding invertebrates prefer leaves conditioned by microorganisms (particularly aquatic hyphomycetes) as they are more palatable and nutritious (Bärlocher, 1985; Maltby, 1992).

9.3.2.5. Statistical and ecological evaluation of concentration–response relationships

Considering the SPGs defined for water organisms (see chapter 5) the statistical analysis of measurement endpoints related to population-level effects are a necessity. Several state-of-the-art techniques are available for univariate analysis (e.g. Williams test; Kruskal–Wallis multiple comparison test; Dunnett's test, Bonferroni U-test) to calculate NOECs and LOECs at the population level. It needs to be noted that more complex micro-/mesocosm experiments may result in considerable inter-replicate variation in measurement endpoints, including population densities (Sanderson et al., 2009). Consequently, to identify the robustness of the LOEC identification the MDD should be reported for single species together with the statistical approach used (see below). Synonyms of MDD are critical boundary (Sokal and Rohlf, 1995), minimum significant difference (Environment Canada, 2005; Van der Hoeven, 2008) and least significant difference (EFSA GMO Panel, 2010). At the population level, concentration–response relationships may also be evaluated by means of logistic or non-linear regression techniques to calculate EC_x values. To ensure that an effect of the PPP is treatment related and not background variability, the EC_x value has to be significant, determined by an adequate statistical test (e.g. the Chi^2-Test for probit or logistic analysis), while also the confidence intervals of the EC_x estimates need to be reported to evaluate the experimental uncertainty associated with the EC_x estimate. Note that population-level NOECs/LOECs and EC_x values usually can be calculated only for taxa that occur in high enough numbers and that dominate the community. It is therefore recommended to also perform univariate statistics on aggregated data, for example total densities of organisms at a higher taxonomic level (e.g. family, order) or on the basis

of densities of organisms with specific ecological traits (see, for example, Liess and Von der Ohe, 2005; McGill et al., 2006; Liess and Beketov, 2011; Gergs et al., 2011).

Some univariate statistical tests (e.g. the Williams test) assume a monotonous concentration–response relationship (increasing effect with increasing concentration). The population and community-level responses observed in a long-term micro-/mesocosm tests, however, may be the result of interplay between direct toxic effects and indirect effects due to shifts in ecological interactions between populations. Factors such as indirect effects may violate the assumption of an increased effect with increasing concentration. For this reason the identification of treatment-related responses should not only be based on statistics but also on ecotoxicological knowledge (to identify the direct toxic effects) and ecological knowledge (to identify possible indirect effects). Statistically significant responses in the same direction (either decreases or increases) on consecutive samplings should be given special weight. Note that because of the high number of possible endpoint–sampling date combinations a statistically significant effect on an isolated sampling may be easily detected while it may be a type II error (accepting the null hypothesis, although the alternative hypothesis is the true state of nature).

Micro-/mesocosm tests also allow the study of treatment-related responses at the community level. To evaluate community-level effects multivariate techniques (e.g. redundancy analysis (RDA) and principle response curves (PRC) in combination with Monte Carlo permutation tests) may be an appropriate tool (e.g. Van Wijngaarden et al., 1995; Van den Brink and Ter Braak, 1998, 1999). An advantage of these multivariate techniques is that they also provide species scores that can be used to identify the most important species/populations explaining the community response. In addition, similarity indices and approaches to describe the treatment-related response of biological communities in terms of traits (e.g. Liess and Beketov, 2011; Van den Brink and Ter Braak, 2012) may be used, e.g. in order to reduce inter-replicate variability. Some approaches require a priori assumptions on sensitivities and traits. A review of the NOEC/LOEC of the most sensitive population in the micro-/mesocosm test system with the NOEC/LOEC values on the basis of trait-based groupings, including different types of PPPs (insecticides, fungicides, herbicides), is a topic for future research. These trait-based groupings may be analysed on the basis of univariate and multivariate techniques, the SPEAR approach, or a combination of these techniques. We foresee that in the future further guidance can be provided on the respective advantages of the various methods.

Currently NOECs/LOECs for measurement endpoints are assessed in the majority of micro-/mesocosm experiments (a limited number of model ecosystem experiments focus on deriving EC_x values), but the statistical power to define these endpoints are routinely not reported in micro-/mesocosm reports present in PPP dossiers. Thus also NOECs may be reported for populations (or aggregated species groups) which are present in the micro-/mesocosm test system, but for which, due to low abundance and variability between replicates, statistically detecting any effects will be impossible. So a study can have a long species/taxa list, but for a limited number of these species/taxa concentration–response relationships can be evaluated. Also the demonstration of recovery depends on the statistical power. In some cases recovery could be demonstrated just by higher variability in the controls and treatments during the recovery period shifting the statistical power in a way that no effects can be demonstrated. Calculating the MDD allows reporting the actual effect which could be determined in the experiment for a given endpoint at a given time. For applying the MDD concept to micro-/mesocosm experiments it is noteworthy that the MDD is particularly important if no effect is observed, since when a LOEC can be calculated the statistical power apparently is high enough to detect an effect. Additionally it is noteworthy that a high MDD for several measurement endpoints is a common phenomenon in micro-/mesocosm studies (since only a limited number of populations dominate the community) but this need not be a reason to reject the study if for several relevant endpoints/populations (e.g. 8 populations of the sensitive taxonomic group) a statistical evaluation can be performed. We recommend that in a first step the MDD should be reported together with the NOEC table for each investigated endpoint in time. We propose clustering the MDDs into five groups and the NOECs in the NOEC table might be shaded (or marked) according to the example shown below (Table 31). It has to be noted that the selection of MDD classes is more or less arbitrary and they

should be revised in future on an appropriate database. In the insecticide case study (see Appendix H) an example is given of how MDDs can be reported.

Table 31: Proposal on classes of minimal detectable differences (MDD) due to treatment-related declines in abundance/biomass (see also Appendix F)

Class	MDD	Comment
0	> 100 %	No effects can be determined
I	90–100 %	Only large effects can be determined
II	70–90 %	Large to medium effects can be determined
III	50–70 %	Medium effects can be determined
IV	< 50 %	Small effects can be determined

The use of MDD is introduced to increase transparency and reliability in endpoints derived from micro-/mesocosm studies. The MDD of critical endpoints should ideally exceed class II. Considering the high level of biological variance in micro-/mesocosms (and natural edge-of-field surface waters), endpoints with lower MDD classes (I–II) may, however, be considered relevant. For a proper evaluation information is also required on the normal fluctuations in population densities (for the organisms of concern) in natural edge-of-field surface waters. It is anticipated that in the coming years more practical experience will be obtained in applying MDDs to evaluate results of micro-/mesocosm experiments. This practical experience is required before more detailed guidance on MDD and the interpretation of micro-/mesocosm endpoints can be provided. The PPR Panel advises the preparation of a specific opinion on the use of MDD and the evaluation of micro-/mesocosm studies.

9.3.3. Interpreting micro-/mesocosm experiments

In recent years, discussions shifted towards the awareness of inconsistencies in both the way the same mesocosm data are interpreted and the AF applied by regulatory experts in different EU Member States. The Dutch Platform for Assessment of Higher-Tier Studies has produced a GD on how micro-/mesocosm data should be presented and evaluated in a uniform and transparent manner (De Jong et al., 2008). We propose to largely use this document to present and evaluate micro-/mesocosm studies for regulatory purposes when placing PPPs on the European market. The main aspects to consider and some deviations are presented below.

9.3.3.1. Evaluation of the scientific reliability of the micro-/mesocosms test for PPP authorisation

On the basis of the information presented in section 9.3.2 the following questions should be answered in the evaluation of the scientific reliability of the micro-/mesocosm experiment.

1. *Is the test system adequate and does the test system represent a realistic freshwater community?* (Trophic levels; taxa richness and abundance of (key and sensitive) species; representativeness of the biological traits with respect to vulnerability).

2. *Is the description of the experimental set-up adequate and unambiguous?* (ANOVA or regression design; overall characterisation of the experimental ecosystem/community simulated; measurement endpoints; sampling frequency; sampling techniques).

3. *Is the exposure regime adequately described?* (Method of application of the test substance; relevance for predicted exposure profile in the field; concentration in the application solution; dynamics in exposure concentrations in relevant compartments (e.g. water, sediment); detection limits).

4. *Are the investigated endpoints sensitive and in accordance with the working mechanisms of the compound, and with the results of the first-tier studies?* (Compare selected measurement endpoints with the species potentially at risk as indicated by the lower tiers).

5. *Is it possible to evaluate the observed effects statistically and ecologically?* (Univariate and multivariate techniques applied; unambiguous concentration–response relationships; statistical power of the test; ecological relevance of the statistical output).

The above-mentioned questions could be answered with Yes, Unclear or No, and the answers should be substantiated with arguments.

A further detailed checklist to assess the scientific reliability of the study is given in the table below, followed by another table explaining the reliability index that might be used to classify the overall quality of the study.

Table 32: Checklist for evaluating micro-/mesocosm studies for regulatory purposes (adapted from De Jong et al., 2008)

Items	Notes	Reliability index 1–3 (see Table 33)
Methodology and test description		
1. Substance	**Properly characterised and reported?**	
1.1 Concentration	Identity and amount of a.s. per litre test water?	
1.2 Formulation and purity	Substances in the formulation influencing the working action of the a.s. should be reported	
1.3 Vehicle	In case a vehicle—other than in the formulation—is used, identity and concentration?	
1.4 Chemical analyses	Method, LOQ, LOD, recovery	
1.5 Properties	Relevant for potential fate and effects in test system	
2. Test site, duration	**Properly characterised and reported?**	
2.1 Location	Necessary to make a link between the effects and local environmental conditions, representativeness	
2.2 Test date/duration	Application dates and experimental period?	
2.4 General climatic conditions	Necessary to make a link between the effects and local climatic conditions	
3. Application	**Properly characterised and reported?**	
3.1 Mode of application	Exposure route; spraying or homogenising the a.s. into the test medium?	
3.2 Dosage and exposure	Actual concentrations during the test? Chemical analysis of dosing solution?	
3.3 Application scheme	Necessary to make a link between the test and the intended use of the PPP	
3.4 Conditions during application	Weather conditions during application, wind speed and temperature?	
4. Test design	**Properly designed and reported?**	
4.1 Type and size	e.g. outdoor microcosm, outdoor pond or mesocosm; dimensions	
4.2 Pre-treatment	Proper equilibration?	
4.3 Treatment period	Number and spacing of treatments?	
4.3 Post-treatment	Period long enough to allow expression of effects and recovery?	
4.4 Untreated control	Sufficient number; solvent applied?	
4.5 Replications	Sufficient replications for proper statistical analysis?	
4.6 Statistics	Univariate and multivariate techniques applied	
4.8 Dose–response	Number of test concentrations for finding a dose–response relation (controls excl.)	
4.9 Quality assurance	Study conducted under GLP?	
5. Biological system	**Representative and properly reported?**	

Items	Notes	Reliability index 1–3 (see Table 33)
5.1 Populations	Enough sensitive/vulnerable species of the relevant taxonomic group?	
5.2 Community	The community/ecosystem representative and complete?	
6. Sampling	**Is sampling adequate for risk assessment?**	
6.1 General features	Relevance selected measurement endpoints	
6.2 Actual concentration	Actual concentrations measured in medium and other compartments or biota?	
6.3 Biological sampling	Appropriate methods and frequency?	
Results		
7. Endpoints	**Properly reported?**	
7.1 Type	Reported endpoints relevant for objective of study?	
7.2 Value	Are measured data consistently presented?	
7.3 Verification of endpoint	Test results are verifiable and source data reported	
8. Elaboration of results	**Are conclusions based on measured data? Methodology correct?**	
8.1 Statistical comparison	Data meet requirements for method used?	
8.2 Dose–effect relationship	Minimal detectable difference; consistence of response	
8.3 Population-level responses	Sufficiently reported?	
8.3 Community-level responses	Sufficiently reported?	
9. Control		
9.1 Untreated control	Unexpected effects or disappearance of species?	
9.2 Solvent control	Possible effects caused by solvent?	
10. Classification of effects	**Properly derivable?**	
11. Biological meaning of statistically significant differences	**Sufficiently explained?**	

Table 33: Definition of the three values of the reliability index

Reliability index (Ri)	Definition	Description
1	Reliable	All data are reported, the methodology and the description are in accordance with internationally accepted test guidelines and/or the instructions, all other requirements fulfilled
2	Less reliable	Not all data reported, the methodology and/or the description are slightly deviating from internationally accepted test guidelines or the instructions, without motivation, or not all other requirements fulfilled
3	Not reliable	Essential data missing, the methodology and/or the description are not in accordance with internationally accepted test guidelines and/or the instructions without motivation, or not reported, or important other requirements are not fulfilled

Based on the questions and checklist above an overall reliability index should be assigned to the micro-/mesocosm study. Both reliability index Ri1 and Ri2 tests might be used in the RA, but it may be decided to apply a larger AF in the derivation of the micro-/mesocosm RAC when only an Ri2 study is available on the basis of the most relevant (sensitive or vulnerable) population or community endpoint (see section 9.3.5).

When the micro-/mesocosm study is deemed reliable to use in the effect assessment of the PPP under evaluation the concentration–response relationships should be evaluated. Below effect classes to summarise the concentration–response relationships of micro-/mesocosm experiments are given, based on the definition by Brock et al. (2006) and De Jong et al. (2008) and modified to add the additional information about the MDD (see section 9.3.2.5).

Effect class 0 (*Treatment related effects cannot be evaluated. If this class is consistently assigned to endpoints that are deemed most relevant for the interpretation of the study the regulatory reliability of the micro-/mesocosm tests is questionable*)

Due to e.g. low abundance and variability the MDD was always larger than 100 % so even very strong effects could not be determined for the endpoint evaluated.

Effect class 1 (*No treatment-related effects demonstrated for the most sensitive endpoints*)

No (statistically and/or ecologically significant) effects observed as a result of the treatment. Observed differences between treatment and controls show no clear causal relationship.

Effect class 2 (*Slight effects*)

Effects concern short-term and quantitatively restricted responses usually observed at individual samplings only.

Effect class 3A (*Pronounced short-term effects (< 8 weeks, followed by recovery*)

Clear response of endpoint, but full recovery of affected endpoint within 8 weeks after the first application or, in the case of delayed responses and repeated applications, the duration of the effect period is less than 8 weeks and followed by full recovery [33]. Treatment-related effects demonstrated on consecutive samplings. Note that recovery can only be considered if the MDDs during the recovery period were at least smaller than 100 %. If this is not the case an appropriate higher class has to be selected.

Effect class 3B (*Pronounced effects and recovery within 8 weeks post last application*)

Clear response of the endpoint in micro-/mesocosm experiment repeatedly treated with the test substance and that lasts longer than eight weeks (responses already start in treatment period), but full recovery of affected endpoint within eight weeks post last application. Note that recovery can only be considered if the MDDs during the recovery period were at least smaller than 100 %. If this is not the case, an appropriate higher class has to be selected.

Effect class 4 (*Pronounced effect in short-term study*)

Clear effects (e.g. large reductions in densities of the population) observed, but the study is too short to demonstrate complete recovery within eight weeks after the (last) application.

Effect class 5A (*Pronounced long-term effect followed by recovery*)

Clear response of sensitive endpoint, effect period longer than 8 weeks and recovery did not yet occur within 8 weeks after the last application but full recovery is demonstrated to occur in the year of application. Note that recovery can only be considered if the MDDs during the recovery period were at least smaller than 100 %. If this is not the case an appropriate higher class has to be selected.

Effect class 5B (*Pronounced long-term effects without recovery*).

[33] An endpoint is considered as recovered if the MDD allows statistical evaluation during the relevant recovery period (so excluding MDD class 0) and the conclusion of no statically significant effect between treated systems and controls is not caused by a decline of that endpoint in controls (e.g. at the end of the growing season). If these criteria are violated a higher effect class has to be selected.

Clear response of sensitive endpoints (> 8 weeks post last application) and full recovery cannot be demonstrated before termination of the experiment or before the start of the winter period.

9.3.4. Variability in concentration–response patterns between micro-/mesocosm experiments exposed to the same PPP

9.3.4.1. Short-term pulsed exposure

For the interpretation of micro-/mesocosm experiments an important question at stake is whether concentration–response relationships are reproducible. Effect classes 1 and 2 concentration of the most sensitive measurement endpoints in the micro-/mesocosm experiment may be used as estimates of the ecological threshold concentrations of PPPs (not considering ecological recovery) while an effect class 3A concentration of the most sensitive measurement endpoints, may be used as the study-specific NOEAEC (no observed ecologically adverse effect concentration), an estimate that may take into account ecological recovery. For only a few test substances more than two appropriately preformed micro-/mesocosm experiments are available when considering the criteria described in section 9.3.3 and a similar exposure regime. The data presented in Appendix E for the insecticides chlorpyrifos, lambda-cyhalothrin and esfenvalerate are representative for (short-term) pulsed exposure regimes. It seems that for these insecticides effect class 1–2 concentrations of the most sensitive measurement endpoints derived from different micro-/mesocosm experiments show lower variability than higher effect classes. Note that the chlorpyrifos, lambda-cyhalothrin and esfenvalerate studies comprised test systems that considerably varied in dimensions and complexity of community structure (plankton-dominated, macrophyte-dominated, lentic, lotic), but always contained several dominant populations of arthropods (but not always well-established populations of insects). Nevertheless, the comparison of micro-/mesocosm experiments performed with these insecticides suggests that a small AF may be sufficient to extrapolate effect class 1 and/or effect class 2 concentration of the most sensitive measurement endpoints derived from a well-performed micro-/mesocosm study with a well-defined exposure regime. In addition, for the same PPP and a similar exposure regime these effect class 1 and effect class 2 concentrations do not overlap with the range of concentrations for higher effect classes (effect classes 3–5).

If an effect class 3A concentration (of most sensitive measurement endpoints) for short-term exposures is considered acceptable, it appears from the data presented in Appendix E that for chlorpyrifos and lambda-cyhalothrin an AF of 3 may be necessary to cover the variability in observed concentration–response patterns that include ecological recovery, if a single high-quality micro-/mesocosm experiment is available. Also in the case of esfenvalerate, applying an AF of 3 to the effect class 3A concentration overall avoids the occurrence of unacceptable class 4–5 effects caused by pulsed exposures in hydrologically closed systems (lentic micro-/mesocosms or recirculating experimental streams) (Tables E1.1 to E1.3 in Appendix E).

In accordance with the data for chlorpyrifos, lambda-cyhalothrin and esfenvalerate described in Appendix E, in lake enclosure studies exploring effects of a single application of pentachlorophenol to plankton communities in spring, summer, autumn and winter, a low variability in threshold levels for effects (effect class 1 concentrations based on peak exposure) was observed. In these lake enclosure experiments ($n = 4$) the variability in ecological threshold concentration varied by approximately a factor of 2 (Willis et al., 2004).

9.3.4.2. Long-term exposure to the same PPP

Again, considering the criteria mentioned in section 9.3, for a few PPPs only three or more appropriate micro-/mesocosm studies are available mimicking a more or less constant chronic exposure regime. In lentic test systems the treatment-related responses caused by a long-term chronic exposure regime to the fungicide carbendazim resulted in similar effect class 1 concentrations, suggesting little variability in threshold levels for effects between studies (Table E1.4 in Appendix E). However, long-term exposure studies with the herbicide atrazine (Table E1.5 in Appendix E) revealed a considerable

overlap between effect class 1 and effect class 2 concentrations of the most sensitive measurement endpoints. In addition, an overlap between effect class 2 and effect class 3–5 concentrations was observed as well for atrazine. As explained in Appendix E, differences in concentration–response patterns between studies performed with the photosynthesis-inhibiting herbicide atrazine might be explained by differences in light conditions between indoor and outdoor studies presented in Table E1.5. Nevertheless, if we consider the atrazine data representative for chronic exposure regimes of other photosynthesis-inhibiting herbicides, and from a regulatory point of view an effect class 2 response is acceptable as an estimate that approaches the threshold level of effects, an AF of 2 –3 seems to be necessary to address the variability in concentration–response patterns between well-performed model ecosystem experiments mimicking a chronic exposure regime. Applying an AF of 2–3 to effect class 2 concentrations presented in Table E.5 (see Appendix E) will, with a high probability, avoid unacceptable class 3 to 5 effects caused by long-term exposure.

9.3.5. How to derive a RAC from an appropriate micro-/mesocosm experiment and how to link it to PEC

Figure 10 and Table 34–Table 37, present proposals for the derivation of the RACs within acute ($RAC_{sw;ac}$) and chronic ($RAC_{sw;ch}$) effect assessment schemes on the basis of appropriate micro-/mesocosm experiments. A distinction is made in RACs derived on the basis of the ETO (ETO-RAC) and ERO (ERO-RAC). Note that when it is possible to derive an ERO-RAC it is always advisable to explore the possibility to also derive an ETO-RAC from the same micro-/mesocosm study.

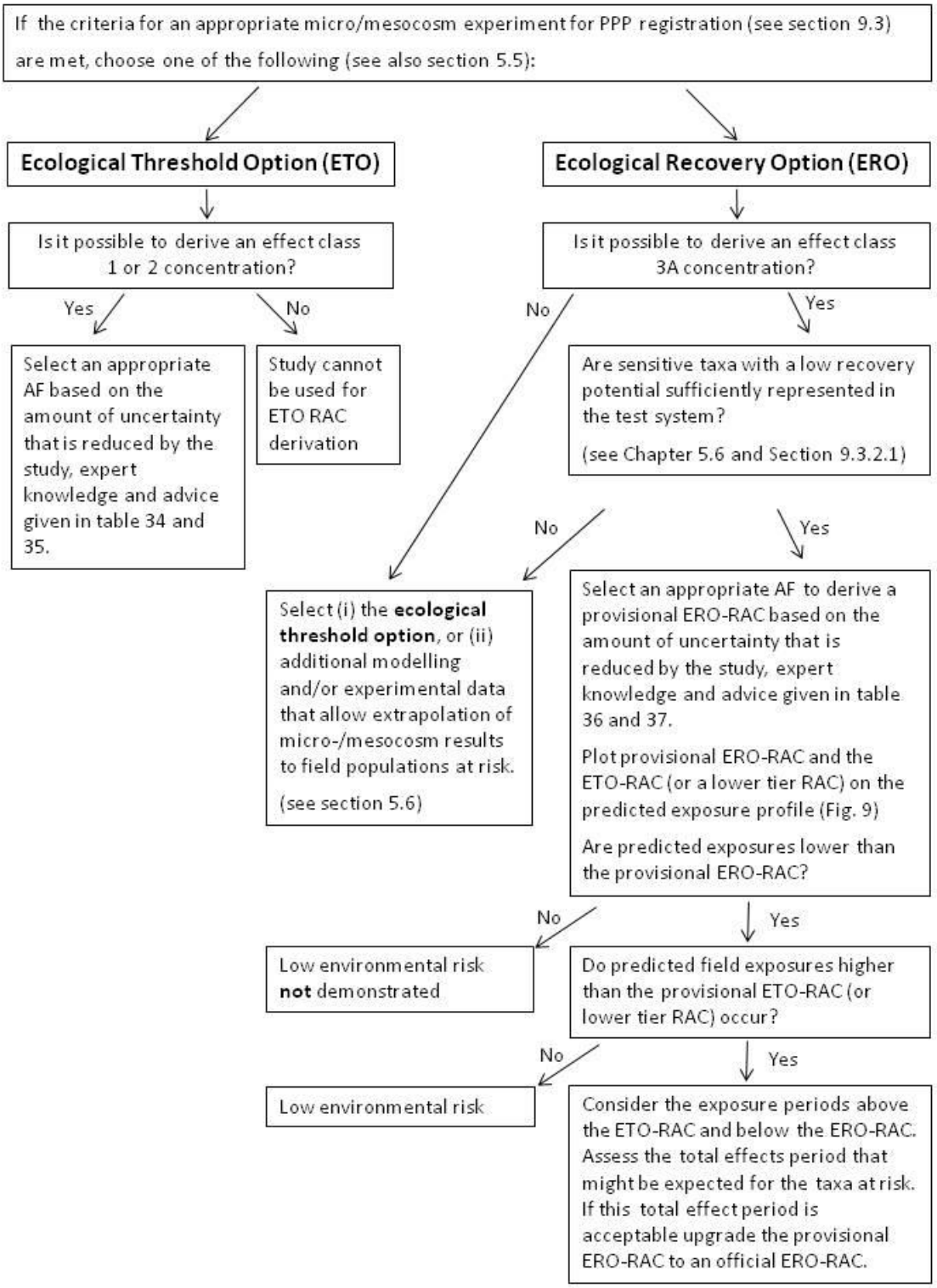


Figure 10: Decision scheme for the derivation of RACs from appropriate micro-/mesocosm experiments on the basis of the ETO (ETO-RAC) or ERO (ERO-RAC)

9.3.5.1. Selecting and extrapolating micro-/mesocosm results

For RAC derivation only those micro-/mesocosm studies should be selected that largely meet the criteria mentioned in sections 9.3.2 and 9.3.3. The RAC may be derived by applying an AF (for spatio-temporal extrapolation) to the study-specific NOEC (ETO) or the NOEAEC that takes into account ecological recovery (ERO). The size of AF should address the model ecosystem to field ecosystem

extrapolation. Edge-of-field surface waters in Europe show a large variability in ecosystem structure and functioning and a specific microcosm or mesocosm experiment mimics only one of the possible field assemblages. Within this context it is worthwhile mentioning that most insight in the variability in concentration–response relationships for PPPs is available for lentic communities in micro-/mesocosms.

Addressing the uncertainty with respect to the model ecosystem–field extrapolation for the threshold level of effects, amongst others (see criteria in section 9.3.3), depends on the relevance of the tested assemblages for the sensitivity (and vulnerability if the recovery option is selected) of species that occur in the type of edge-of-field surface water potentially at risk. In addition, other higher-tier information available (e.g. laboratory toxicity data for additional test species and other micro-/mesocosm experiments) may address this uncertainty. Usually more ecotoxicological data are available for species occurring in lentic edge-of-field surface waters (ponds, drainage ditches) than for species typical for edge-of-field streams. For example, the available toxicity data for insecticides and ETP taxa may be scarce, while these taxonomical groups often are more abundant in lotic than in lentic surface waters. If there are reasons to assume that, for example, a lentic micro-/mesocosm does not sufficiently represent sensitive taxa typical for streams, and these taxa are probably more sensitive than lentic taxa, this may be a reason not to accept the micro-/mesocosm study for RAC derivation to be used in the RA of predicted exposures in streams. The same type of reasoning may be valid when extrapolating lotic data to ponds or drainage ditches when taxonomic groups are sensitive that usually are more abundant in lentic waters (e.g. floating macrophytes, chironomids, copepods). Note, however, that Maltby et al. (2005) assessed the influence of lentic versus lotic habitat on the species sensitivity distributions of arthropods to eight insecticides and that there was no consistent pattern in the relative sensitivity between lentic and lotic species.

If a micro-/mesocosm is available that does not fully meet the criteria (e.g. less than minimum number of relevant populations or no effect class 1–2 can be derived), but adequate additional information is provided to complement the missing data/information, it may be decided on a case-by-case basis to use these data in RAC derivation. However, this should only be considered when adequate additional information is available, e.g. another micro-/mesocosm study that not fully meets the criteria but provides concentration–response relationships for other sensitive populations, or additional long-term laboratory toxicity studies with sensitive species present in the micro-/mesocosm test system to derive the NOEC for these species. In those cases an AF higher than the proposed ranges may be needed.

9.3.5.2. Peak, nominal or TWA concentrations of RAC and PEC used for risk assessment

If the study is triggered by the tier 1 acute core data and the duration of the pulse exposure in the micro-/mesocosm experiment appears to be shorter than that predicted for the field, the following approach is recommended. Express the treatment-related responses (short-term exposure effect assessment) in terms of the initial 0–48 to 0–96 hours TWA concentration (instead of the measured peak concentration) as measured in the test systems and compare the final $RAC_{sw;ac}$ estimate to the PEC_{max}. The duration of 48–96 hours is selected since in the first-tier acute effect assessment this time frame corresponds with the duration of most standard acute tests. In addition, a similar procedure is proposed to derive a MAC-EQS from a micro-/mesocosm test within the context of the Water Framework Directive. In this case the TWA approach should not be used for the short-term exposure assessment, because then the worst-case assumption of the approach is violated—instead the $PEC_{sw,max}$ should be used. Another promising option is to explore adequate methods to extrapolate concentration–response relationships for shorter pulse exposures to that of broader ones (e.g. the use of TK/TD models for relevant sensitive organisms may be promising); however, no detailed guidance is included here for the time being. This might be updated based on a future activity of the PPR Panel on aquatic effect modelling.

To evaluate chronic risks (triggered by the tier 1 chronic core data) either the peak concentration or a TWA concentration of the PPP in the relevant matrix (water, sediment) may be used as estimate of $RAC_{sw;ch}$ and/or as PEC estimate (see chapter 4). If the TWA approach is appropriate (for criteria, see

chapter 4) and used for the effects assessment, the selected TWA time window should coincide with the application period of the test substance. For both RAC and PEC estimation the selection of the length of the TWA time window should be based on ecotoxicological considerations (e.g. A:C ratio; time-to-onset-of-effect information; length of the most sensitive life stage of the organisms at risk) and guided by the length of the relevant chronic toxicity tests that triggered the micro-/mesocosm experiment. If the TWA approach is considered appropriate (for criteria, see chapter 4) we propose to adopt a default time window of seven days for the TWA estimate of the long-term PEC (= $PEC_{sw;7d\text{-}twa}$), at least if no scientific arguments are provided to shorten or lengthen this default time window. This proposal follows the recommendation of the ELINK workshop. Note that for a worst-case approach the time window for the TWA effect class concentration estimate in the micro-/mesocosm study should not be smaller than the selected TWA time window for the PEC estimate in the field. In addition, in the case of single application studies the exposure should not decline to levels lower than 20 % of nominal during the time-window for the TWA effect class concentration estimate, or, in the case of a relative fast-dissipating substance and repeated weekly applications, the TWA time window should not exceed the application period of the PPP in the micro-/mesocosm study. The application period is the period in which repeated pulse applications occur. When e.g. a seven-day time window is adopted for the PEC, the 'effect class' concentrations derived from a micro-/mesocosm experiment characterised by three weekly treatments can be expressed in terms of a TWA concentration that is ≥ 7 days and ≤ 21 days if in the test systems the PPP is not very persistent.

In case the TWA approach is deemed not to be appropriate in the chronic RA, and consequently the $PEC_{sw;max}$ is used as field exposure estimate, the 'effect class' concentrations derived from a mesocosm experiment simulating long-term exposure may be expressed in terms of the nominal, peak or average concentration measured/calculated during the application period (or the period in which the exposure remains more or less constant in the micro-/mesocosm test). Adopting the nominal or measured/calculated peak concentration is justified only if it can be demonstrated that the exposure profile in the micro-/mesocosm experiment overall is realistic to worst case compared with in the relevant field scenario(s). In that case, and if it was demonstrated that the concentration builds up due to repeated treatments, adopting the nominal concentration during the application period can be considered as a more conservative approach than adopting the measured/predicted peak concentration.

9.3.5.3. Deriving a RAC indicative for the ETO (ETO-RAC)

As already discussed in section 5.5, application of the ERO may be uncertain when assessing risks for individual PPPs for their use in crop protection programmes characterised by intensive PPP use (e.g. simultaneous use of PPPs with similar mode of action in tank mixtures or their repeated use). However, a thorough analysis of PPP usage practices in major crops and in different EU Member States is not yet available. In addition, an accurate analysis of potential multi-stress effects of PPPs in edge-of-field surface waters needs RA tools still under development (e.g. ERA guidance for sediment organisms and mechanistic effect models). Current information with respect to multi-stress of PPPs in edge-of-field surface waters seems to indicate that mixture toxicity cannot be excluded but that it usually is one or a few a.s. only that dominate the mixture in terms of toxic units (for further details and references, see section 12.3).

Under the assumptions that (i) the sensitivity of the assemblages in appropriate micro-/mesocosm tests are representative for those in edge-of-field surface waters, and (ii) that the observed variability in threshold concentrations for effects between micro-/mesocosm tests with the substances chlorpyrifos, lambda-cyhalothrin, esfenvalerate, pentachlorophenol, carbendazim and atrazine (see section 9.3.4) is generally valid for PPPs, the AF to address variability in effect class 1–2 concentrations between freshwater communities that contain sensitive populations potentially at risk may be small.

Table 34 presents a proposal for the derivation of the $RAC_{sw;ac}$ (triggered by the tier 1 acute core data) addressing the ETO for edge-of-field surface waters on the basis of an appropriate micro-/mesocosm experiment. The AFs presented in the table below are proposed for studies that are appropriately designed to address the risk and uncertainty identified at lower tiers, e.g. studies in which a

sufficiently low MDD was obtained for adequate number of species (see section 9.3.2). If this is not the case, the AF needs to be adjusted.

Table 34: Proposal for the derivation of the $RAC_{sw;ac}$ (triggered by tier 1 acute core data) addressing the ETO on the basis of an appropriate micro-/mesocosm experiment. Note that in the same study several treatment levels may result in effect class 1 responses for sensitive measurement endpoints. In that case the highest treatment level showing an overall effect class 1 response should be selected for ETO-RAC derivation. Alternatively, if in the same study several treatments result in effect class 2 responses in first instance the lowest treatment level showing an overall effect class 2 response should be selected for ETO-RAC derivation. On a case-by-case basis and expert judgement it may be decided to select a higher treatment level as overall effect class 2 concentration

	Assessment factor for ETO-$RAC_{sw;ac}$ derivation (ecological threshold option)	**Field exposure concentration to compare with the $RAC_{sw;ac}$**
Effect class 1 concentration		
Is rate of dissipation of the a.s. in test system realistic to worst case when compared to that predicted for the field? *Yes*: Base effect estimate on nominal or measured peak concentration in test system *No*: Base effect estimate on, for example, the initial 48 h TWA concentration in test system or apply appropriate extrapolation techniques	2[(a)]	$PEC_{sw;max}$
Effect class 2 concentration		
Is rate of dissipation of the a.s. in test system realistic to worst case when compared to that predicted for the field? *Yes*: Base effect estimate on nominal or measured peak concentration in test system *No*: Base effect estimate on, for example, the initial 48 h TWA concentration in test system or apply appropriate extrapolation techniques	2–3[(a)] The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

Table 35 presents a proposal for the derivation of the $RAC_{sw;ch}$ (triggered by the tier 1 chronic core data) addressing the ETO for edge-of-field surface waters on the basis of appropriate micro-/mesocosm experiments.

Table 35: Proposal for the derivation of the $RAC_{sw;ch}$ (triggered by tier 1 chronic core data) addressing the ETO on the basis of an appropriate micro-/mesocosm experiment. Note that in the same study several treatment levels may result in effect class 1 responses for sensitive measurement endpoints. In that case the highest treatment level showing an overall effect class 1 response should be selected for ETO-RAC derivation. Alternatively, if in the same study several treatments result in effect class 2 responses in first instance the lowest treatment level showing an overall effect class 2 response should be selected for ETO-RAC derivation. On a case-by-case basis and expert judgement it may be decided to select a higher treatment level as overall effect class 2 concentration

	Assessment factor for ETO-$RAC_{sw;ch}$ derivation (ecological threshold option)	**Field exposure concentration to compare with the $RAC_{sw;ch}$**
Effect class 1 concentration		
Based on time weighted average concentration in test system during the application period	2[(a)]	$PEC_{sw;max}$ or $PEC_{sw;twa}$. Based on expert judgement by considering the criteria mentioned in chapter 4
Based on nominal or peak concentration in test system if the long-term exposure regime (e.g. due to repeated pulses) is realistic to worst case compared with the predicted field exposure profile	2[(a)]	$PEC_{sw;max}$
Effect class 2 concentration		
Based on TWA concentration in test system during the application period	2–3[(a)]. The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$ or $PEC_{sw;twa}$. Based on expert judgement by considering the criteria mentioned in chapter 4
Based on nominal or peak concentration in test system if the long-term exposure regime (e.g. due to repeated pulses) is realistic to worst case compared to the predicted field exposure profile)	2–3[(a)]. The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. If the available micro-/mesocosm studies are of the same quality, the AF may be applied to the geometric mean value of the effect class 1 or effect class 2 concentrations derived from the different studies. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

In Tables 34 and 35 to extrapolate an effect class 2 concentration an AF in the range 2–3 is proposed. Factors that can be considered to justify the lower AF within the proposed range to derive RACs:

- number of replicates (more than the minimum requirement);

- number of exposure concentrations tested (larger than the minimum requirement (five concentrations);
- sufficient pre-treatment period to allow the community to be well-established in the system;
- ecological relevance and richness of species of the community tested (more than the minimum 8 sensitive/vulnerable species with acceptable MDD);
- worst-case character of the exposure concentrations tested relative to the predicted exposure scenario.

Note that most of the factors mentioned above are addressed when assigning a reliability index (Ri) to the micro-/mesocosm study. Studies assigned to Ri1 and Ri2 (see section 9.3.3 and Table 33) can be used to derive RACs, and a lower AF within the proposed range can be selected for a Ri1 study than for a Ri2 study.

9.3.5.4. Deriving a RAC on the basis of ERO (ERO-RAC)

Not exceeding the SPGs for aquatic key drivers described in chapter 5, effect class 3A concentrations from appropriate micro-/mesocosm tests might be used to derive a RAC in line with the recovery option.

Again, under the assumptions that (i) the sensitivity of the assemblages in appropriate micro-/mesocosm tests are representative for those in edge-of-field surface waters, and (ii) that the observed variability in threshold concentrations for effects between micro-/mesocosm tests with the substances chlorpyrifos, lambda-cyhalothrin, esfenvalerate, pentachlorophenol, carbendazim and atrazine (see section 9.3.4) is generally valid for PPPs, the AF to address the extrapolation of an effect class 3A concentration from a micro-/mesocosm test system to the field needs to be larger than that to derive the RAC representative for the ETO (see section 9.3.4).

Table 36 and Table 37 present proposals for respectively the derivation of the $RAC_{sw;ac}$ (triggered by tier 1 acute core data) and the $RAC_{sw;ch}$ (triggered by tier 1 chronic core data) addressing the ERO for edge-of-field surface waters on the basis of an appropriate micro-/mesocosm experiment.

When using an effect class 3A concentration to derive the RAC special attention should be paid to the representativeness of the enclosed community in the test system for potentially sensitive invertebrate populations with a complex uni-/semivoltine life cycle and limited dispersal abilities and/or for potentially sensitive macrophytes with a relatively slow growth rate. If these populations are insufficiently represented in the test system additional information (e.g. additional population studies) and/or extrapolation techniques (e.g. population models) may be required, or simply the RAC indicative for the ecological threshold level (ETO-RAC) needs to be used in the RA.

Table 36: Proposal for the derivation of the $RAC_{sw;ac}$ (triggered by tier 1 acute core data) addressing the ERO on the basis of an appropriate micro-/mesocosm experiment. If in the same study several treatments result in effect class 3A responses for sensitive measurement endpoints in first instance the lowest treatment level showing an overall effect class 3A response should be selected to derive the ERO-RAC. On a case-by-case basis and expert judgement it may be decided to select a higher treatment level as overall effect class 3A concentration

	Assessment factor for ERO-$RAC_{sw;ac}$ derivation (ecological recovery option)	**Field exposure concentration to compare with the $RAC_{sw;ac}$**
Effect class 3A concentration		
Maximum magnitude of temporal effects may be medium to large. *Is rate of dissipation of the a.s. in test system realistic to worst case when compared with that predicted for the field?* *Yes*: Base effect estimate on nominal or measured peak concentration in test system. *No*: Base effect estimate on, for example, the initial 48 h TWA concentration in test system or, apply appropriate extrapolation techniques or, consider the ETO (Table 34)	3–4[(a)]. The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

Table 37: Proposal for the derivation of the $RAC_{sw;ch}$ (triggered by tier 1 chronic core data) addressing the ERO on the basis of an appropriate micro-/mesocosm experiment. If in the same study several treatments result in effect class 3A responses for sensitive measurement endpoints in first instance the lowest treatment level showing an overall effect class 3A response should be selected to derive the ERO-RAC. On a case-by-case basis and expert judgement it may be decided to select a higher treatment level as overall effect class 3A concentration

	Assessment factor for ERO-$RAC_{sw;ch}$ derivation (ecological recovery option)	**Field exposure concentration to compare with the $RAC_{sw,ch}$**
Effect class 3A concentration		
Based on TWA concentration in test system during the application period	3–4[(a)]. The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$ or $PEC_{sw;twa}$. Based on expert judgement by considering the criteria mentioned in chapter 4
Based on nominal or peak concentration in test system if the long-term exposure regime (e.g. due to repeated pulses) is realistic to worst case compared with the predicted field exposure profile	3–4[(a)]. The definitive choice of the AF is based on expert judgement by considering the criteria mentioned in section 9.3 and ecological information on the type of edge-of-field surface water at risk	$PEC_{sw;max}$

(a): If several adequate micro-/mesocosm studies or other adequate higher tier studies (e.g. monitoring, relevant population experiments or modelling) are available the AF should be applied to the RAC derived from the most appropriate micro-/mesocosm study (see criteria section 9.3.5.1) for the specific case, considering a weight of evidence approach. If the available micro-/mesocosm studies are of the same quality, the AF may be applied to the geometric mean value of the effect class 3A concentrations derived from the different studies. Check whether the micro-/mesocosm derived RAC is also protective for vertebrates (e.g. RACs on the basis of tier 1 and tier 2 approaches) and select the lowest value as final RAC.

In Tables 36 and 37 to extrapolate an effect class 3A concentration an AF in the range 3–4 is proposed. Factors that can be considered to justify the lower AF within the proposed range to derive RACs are the same as mentioned above in section 9.3.5.4. Again, a lower AF within the proposed range can be selected for a Ri1 study than for a Ri2 study (see section 9.3.3 and Table 33).

Finally, according to ELINK (Brock et al., 2010a) and the procedure described in section 9.1.5 the derived ERO-RAC from a micro-/mesocosm experiment should in first instance be considered as a 'provisional ERO-RAC'. An appropriate RA can be performed only by plotting this ERO-RAC and the RAC indicative for the ecological threshold level for effects (= ETO-RAC) on the predicted field exposure profile (see Figure 9). If in the appropriate edge-of-field scenario the pulses are lower than the ERO-RAC value based on effect class 3A concentration but higher than the ETO-RAC, the interval between successive pulses should be carefully considered based on the difference in number and spacing of the pulse exposures simulated in the micro-/mesocosm test and the predicted filed exposure profile. If, for example, the number of pulse exposures is larger in the predicted field exposure profile than in the micro-/mesocosm test under evaluation, the total period of possible effects can be estimated and considered in the final RA (whether the recovery option can or cannot be used, Figure 10).

10. Non-testing methods, metabolites, impurities and formulations with more than one active substance

10.1. Non-testing methods

This section provides guidance on the use of non-testing methods in PPP RA, such as (Q)SAR, expert test systems and analogue read-across, as tools for deriving intrinsic properties of chemicals. All of these approaches have a role in extending and supplementing the existing information and hereby minimising the need for additional testing, in particular, to address potential risks for metabolites (section 10.2) and to reduce toxicity tests with vertebrates (see section 11.4). In addition, non-test methods are important tools for prioritising chemicals for further consideration, assessment and/or testing and in the planning of further testing. The development and application of all kinds of non-testing methods is based on the similarity principle, that is, the hypothesis that similar compounds should have similar biological activities and the methods may therefore in some cases provide predictions that are so reliable that they can be used to substitute experimental data for several types of hazard related endpoints, for example, mortality, reproduction or endocrine effects. In this regard it should be noted that a positive prediction made by the application of a non-test method for an effect (e.g. reproductive toxicity) may be accepted for use to avoid further testing while caution should be applied with negative predictions (i.e. lack of effect) since in most cases not all modes-of-action or mechanisms are covered by the existing non-test method.

When using (Q)SARs, it should be remembered that (Q)SARs are models and are therefore inevitably associated with a degree of uncertainty. This uncertainty is caused predominantly by two different reasons: (a) the inherent variability of the input data used to establish and validate the (Q)SAR model and (b) the uncertainty resulting from the fact that a model can only be a partial representation of reality (in other words, it does not generally model all possible modes of action or mechanisms and hence does not represent all types of chemicals). It is noteworthy that these two types of uncertainty are related to the validation and the applicability domain of the (Q)SAR model respectively. Despite these uncertainties, it is also noted that a (Q)SAR is not only an empirical model, but that it is associated with (1) an underlying dataset used to establish and validate the model, (2) a description of the modelled endpoint, (3) the descriptors and the statistical methods used, (4) a characterisation of the applicability domain and (5) any appropriate mechanistic understanding of the model. As a representation of the training dataset for the model, it averages the uncertainty over all chemicals. Thus, if the model makes reliable predictions within its applicability domain, an individual model estimate will be more accurate than an individual measurement obtained by performing the relevant test.

10.1.1. Area of use

Use of alternative '*in silico'* methods such as qualitative or quantitative structure–activity relationship models (Q)SARs or read-across may be used on their own or in combination with expert systems (see below) as non-testing methods to provide valid endpoints for assessment of hazard and risk. However, often, more robust estimates can be generated by using weight-of-evidence approaches where all available information is taken into account. This could include a combination of (Q)SAR predictions for the same endpoint by different model systems combined with read-across and other available information like non-standard testing data and toxicodynamic/toxicokinetic information from mammals (OECD, 2007, 2009; NAFTA, 2012[34]).

Data provided by non-testing methods shall not be used to substitute experimental data necessary to fulfil the data requirements (Commission Regulation (EU) 283/2013 and 284/2013). However, there may be situations where non-testing methods can be used to address needs for information, rather than deriving new experimental data. Situations where non-testing methods may be used on a more regular

[34] http://www.epa.gov/oppfead1/international/naftatwg/

basis are for metabolites without the toxophore (see section 10.2.7) and for impurities[35]. In addition, (Q)SARs might, together with available test data, be used to rank species for identifying the most likely sensitive taxonomic group to focus experimental testing on (EFSA PPR Panel, 2012a).

10.1.2. Guidance on (Q)SAR

The guidance provided here on the use of (Q)SAR is specified to the use in relation to PPPs. Guidance on the validity of (Q)SAR models and reliability and adequacy of (Q)SAR predictions in general can be found in the ECHA report 'Guidance on information requirements and chemical safety assessment Chapter R.6: (Q)SARs and grouping of chemicals' (ECHA, 2008). Only a short summary of the ECHA guidance is given below on

- how to establish the validity of (Q)SAR models and how to assess the reliability and adequacy of (Q)SAR predictions;
- how to document and justify the use of a (Q)SAR model and where do find information on (Q)SAR models.

More specific guidance on the use of (Q)SAR in PPP RA can also be found in the recently published GD by the North American Free Trade Agreement[36] (NAFTA) Technical Working Group on PPPs (TWG) (NAFTA, 2012).

It is noted that the field of computational toxicity (including (Q)SAR) is rapidly developing, and experience in the regulatory use of computational approaches (including their reporting) is increasing. This guidance document should be considered as a step in a continuously evolving process.

Reporting of validity assessment should follow the OECD (Q)SAR validation format '(Q)SAR Model Reporting Format' (QMRF). Likewise, the specific prediction should be reported in a '(Q)SAR Prediction Reporting Format' (QPRF) (see ECHA, 2008).

10.1.2.1. Model validity

(Quantitative) Structure–activity relationship models are not formally validated as is the case for OECD test guidelines. Instead the OECD has established five internationally agreed principles that can be used to assess the validity of a (Q)SAR model prediction for a given purpose (OECD, 2007):

Principle 1. A (Q)SAR model should be associated with a defined (measurable) endpoint and the related experimental protocols.

Principle 2. An unambiguous algorithm is to ensure transparency in the model algorithm that generates predictions of an endpoint from information on chemical structure and/or physicochemical properties.

Principle 3. A defined chemical domain of applicability of a (Q)SAR model, for which reliable predictions can be generated.

Principle 4. Appropriate measures of goodness-of-fit, robustness (internal performance) and predictivity (determined by external validation).

Principle 5. A mechanistic interpretation of the descriptors used in a model and the endpoint being predicted (if possible).

[35] Any component other than the pure active substance and/or variant which is present in the technical material (including components originating from the manufacturing process or from degradation during storage) [Art. 3 (33) of Regulation (EC) No 1107/2009].

[36] Formalized framework for partnership between United States Environmental Protection Agency Office of Pesticide Programs (US EPA OPP) and the Pest Management Regulatory Agency (PMRA) of Health Canada to develop common approaches to Integrated Approaches to Testing and Assessment (IATA) for the human health and ecological RA of pesticides.

For detailed explanation of each principle see also Annex A of the OECD guidance on (Q)SARs (2007).

10.1.2.2. Reliability and adequacy of (Q)SAR prediction

The determination of whether a (Q)SAR result may be used to replace experimental testing can be broken down into three main steps:

1. an evaluation of the scientific validity (relevance and reliability) of the model;

2. an assessment of the applicability of the model to the chemical of interest and the reliability of the individual model prediction;

3. an assessment of the adequacy of the information for making the regulatory decision, including an assessment of *completeness*, that is, whether the information is sufficient to make the regulatory decision, and if not, what additional (experimental) information is needed.

To be used as a full replacement of an experimental test, all three conditions need to be fulfilled. In cases where some information elements are missing, (Q)SAR results may still be used in the context of a weight of evidence approach.

Step 1

An assessment of the model relevance and reliability follows the five principles mentioned above. It is noted that there is no unique measure of model reliability. In general, model reliability should be regarded as a relative concept, depending on the context in which the model is applied. In other words, a greater or lesser degree of reliability may be sufficient for a given regulatory application.

When evaluating the performance (fitting and external prediction) of (Q)SAR models, several validation parameters exist, for example, predictive squared correlation coefficient (Q^2) (Shi et al., 2001; Schüürmann et al., 2008; Consonni et al., 2009) and average correlation coefficient (r^2). Recently, the concordance correlation coefficient (CCC) proposed by Chirico and Gramatica (2011) has been compared with commonly acceptance thresholds ($Q^2 = 0.6$, average $r^2 = 0.5$). The CCC (Lin, 1989) is similar to the correlation coefficient (linear alignment), but, in addition, it takes into account the closeness to the diagonal (perfect match). A CCC threshold value of 0.85 has been claimed to be the most restrictive in the acceptance of (Q)SAR model estimates (Chirico and Gramatica, 2012). This indicates that CCC is a validation parameter in a precautionary approach for assessing accurate predictions. It is noted, however, that any validation should always include visual inspection of the experimental versus predicted plot, in order to not overlook biases in data set (e.g. location shift and scale shift).

Step 2

Assessment of model validity is a necessary but not sufficient step in assessing the acceptability of a (Q)SAR result. Assuming that the model is considered valid, the second and crucial step is to evaluate the reliability of prediction for a specific compound. There are four main questions to address. (i) Is the chemical of interest within the scope of the model, according to the defined applicability domain of the model? (ii) Is the defined applicability domain suitable for the regulatory purpose? (iii) How well does the model predict chemicals that are *similar* to the chemical of interest? (iv) Is the model estimate reasonable, taking into account other information?

Step 3

Experience with the use of (Q)SAR data in a regulatory context is relatively limited compared with acceptance of test data (including data on laboratory animals). In a regulatory context, experience in the regulatory use of non-testing data have often been obtained by following a learning-by-doing

approach[37], with the learning being documented in draft assessment reports and/or guidelines/background documents for the particular regulatory area[38]. Only limited guidance on the acceptance of (Q)SARs can be given at this moment. However, three important principles could be outlined here, as already explained in the TAPIR report (ECB, 2005): (i) the principle of proportionality expresses the relationship between the amount of information needed and the severity of the decision; (ii) the principle of caution (or conservativeness) expresses the relationship between the amount of information needed and the (likely) consequence(s) of the decision based on that information being wrong; (iii) the level of confidence and precision of a non-testing prediction is higher when the predicted value is close to a regulatory cut off value/decision point that when it is clearly far away for the cut off. It is noted that the same applies for test data and that this issue is not dependent of the state of science but rather a consequence of human decision making systems.

A consequence of these principles, which also applies to test data, is that the relationship between scientific validity and the reliability of information, and hence its regulatory acceptability, should not be regarded as a constant relationship, but a relationship which varies according to the decision being made and the particular circumstances involved in individual cases.

A formal adoption of (Q)SAR models or other non-testing methods are not foreseen under the registration, evaluation, authorisation and restriction of chemical substances (REACH) (i.e. no official, legally binding list of (Q)SAR methods). Instead, acceptance under REACH will involve initial acceptance by industry and subsequent evaluation by the authorities, on a case-by-case basis. The same will be the case for PPP registration. Use of (Q)SARs in such regulatory contexts is also a learning process, therefore, the PPR Panel recommends interaction between relevant EU agencies in order to come to a common approach.

10.1.3. Available (Q)SAR methods, expert systems and read-across

Many (Q)SAR models that estimate the toxicity to aquatic organisms are available, for example, ECOSAR (US EPA),[39] (Q)SAR Application Toolbox (OECD), The Danish (Q)SAR Database,[40] DEMETRA (EU),[41] TOPKAT, ChemProp[42] and/or expert systems proposed (e.g. Escher et al. (2006)). A (Q)SAR model inventory is available at the homepage of the EC Joint Research Centre.[43]

10.1.3.1. ECOSAR

ECOSAR (Ecological Structure Activity Relationships) is a freely available software system (US EPA, 2008) which matches the structure of a query organic substance to one (or more) of its defined chemical class(es). For most classes, aquatic ecotoxicity values are predicted using available linear correlations between toxicity and hydrophobicity. ECOSAR predicts acute (short-term) toxicity and chronic (long-term or delayed) toxicity to aquatic organisms such as fish, aquatic invertebrates and aquatic plants. If not available experimentally, P_{ow} is estimated for the query molecule using KOWWIN. In 2012, ECOSAR included 111 organic chemical classes in three main groups, of which the group 'organic chemicals with excess toxicity' contains several PPP classes (e.g. carbamates, imidazoles, neonicotinoids, pyrethroids, sulfonylureas, triazines).

[37] E.g. US or OECD High Production Volume Chemicals (HPVC) programmes and to a lesser extent in the former EU ESR Programme on existing industrial HPVC.
[38] Experience should be compiled in the next update of this Aquatic Guidance Document.
[39] ECOSAR (Ecological Structure Activity Relationships) is a freely available piece of software that can be downloaded from the US EPA website (US EPA 2008a).
[40] A database containing (Q)SAR predictions for the environment and human health for 180.000 chemicals (http://130.226.165.14/index.html).
[41] This software tool for the ecotoxicity prediction of pesticides and metabolites was developed as part of an EC funded project named DEMETRA (http://www.DEMETRA-tox.net).
[42] Chemical Properties Estimation Software System (ChemProp) 5.2.7, 2012. UFZ Department of Ecological Chemistry, http://www.ufz.de/index.php?en=6738
[43] http://qsardb.jrc.ec.europa.eu/qmrf/

10.1.3.2. OECD (Q)SAR Application Toolbox

The Toolbox[44] is a software application intended to be used to fill gaps in toxicity and ecotoxicity data, which are needed for assessing the hazards of chemicals. The Toolbox incorporates databases on chemical data (e.g. properties), experimental toxicological and ecotoxicological data and estimated values from a large range of QSAR tools, together with incorporated QSAR modelling and expert systems, built within a regulatory application chassis. This package therefore allows the user to perform a number of functions:

- Identify analogues for a chemical, retrieve experimental results available for those analogues and fill data gaps by read-across or trend analysis;
- Categorise large inventories of chemicals according to intrinsic chemical properties ('profilers') related to, for example, physical chemical properties, chemical reactivity related to various mechanisms or modes-of-action;
- Functionalities for assessment of metabolites of chemicals even though it does not also contain probability estimates or toxicokinetic information/predictions of those metabolites.
- Fill data gaps for any chemical by using the library of QSAR models;
- Evaluate the robustness of a potential analogue for read-across;
- Evaluate the appropriateness of a (Q)SAR model for filling a data gap for a particular target chemical; and
- Build QSAR models;
- Functionalities by which documentation of the performed analysis can be provided (combination of automated reporting which can be manually improved/detailed).

10.1.3.3. The Danish (Q)SAR database

This database is freely available on the Internet and contains (Q)SAR predictions from over 70 (Q)SAR models for approximately 180 000 chemicals. The (Q)SAR models encompass endpoints for physicochemical properties, fate, ecotoxicity, absorption, metabolism and toxicity. The majority of estimates are from models developed for mammalian (human) toxicity endpoints, in particular MULTICASE (Multiple Computer Automated Structure Evaluation, Multicase Inc, Ohio, USA). These predictions also contain simple statements in respect to whether or not the individual prediction is within the structural applicability domain of the model (yes/no). Estimates for many of the environmental properties come from the Epiwin software developed by Syracuse Research Co-operation on behalf of the US Environmental Protection Agency (EPA). However, the acute toxicity models for the environment derive from models developed by the Danish EPA. Estimates from a few literature-based models are also included.

10.1.3.4. DEMETRA

This software tool for the ecotoxicity prediction of PPPs and metabolites was developed as part of an EC-funded project named DEMETRA (http://www.DEMETRA-tox.net). The programme allows the prediction of PPP toxicity in fish, daphnids, bees and quail (oral and dietary exposure) and incorporates predictive models for five specific endpoints, with each hybrid combinative model incorporating an intelligent integration of several individual validated (Q)SAR models. The models and software were developed with the aim of regulatory use and developed according to strict quality criteria according to OECD guidelines, using only experimental data produced according to official guidelines and validating using external test sets. The models are applicable to PPPs (and metabolites/impurities) in general and not specific chemical classes. The predictive models within

[44] The most recent OECD (Q)SAR Application Toolbox (version 3) was launched in October 2012 at: http://www.oecd.org/env/chemicalsafetyandbiosafety/assessmentofchemicals/theoecdqsartoolbox.htm#Download_qsar_application_toolbox

DEMETRA are a hybrid combination of two or more individual models, therefore minimum and maximum values are also computed based on the minimum or maximum predicted values of the individual models (these values do not refer to the range of the hybrid model).

10.1.3.5. TOPKAT

TOPKAT contains a range of cross validated (Q)SARs, which are multivariate statistical relationships between experimentally derived toxicity data and chemical descriptors that quantify chemical transport properties and biochemical interaction with the target site. It also provides the user with a measure of whether the query molecule fits within the prediction space of the chosen relationship and therefore whether the estimation is reliable.

10.1.3.6. ChemProp

The chemical properties estimation software, ChemPropx1, was developed at the Helmholtz Centre for Environmental Research (UFZ) in Leipzig more than 15 years ago (Schüürmann et al., 1997), and became available free-of-charge with completion of the EU-funded OSIRISx2 project. ChemProp covers a wide range of *in silico* methods such as qualitative and quantitative structure–activity relationships (QSARs), computerised read-across, structural alerts, increment methods based on structural fragments, and linear solvation–energy relationships (LSERs) for predicting fate-related partition coefficients and physicochemical properties, physiologically based pharmacokinetic models (PBPK) - relevant to partition coefficients, environmental half-lives, ecotoxicological endpoints, and human toxicological endpoints. Particular features include, among others, fully computerised atom-centred fragment (ACF)-based read-across schemes for predicting aquatic toxicity (Kühne et al., 2013; Schüürmann et al., 2011).

10.1.3.7. Approach of Escher et al.

The approach proposed by Escher et al. (2006) uses the principle of the toxic ratio (TR) (Verhaar et al., 1992) of the parent PPP to estimate the maximum potency of a metabolite. The TR is the ratio between baseline toxicity, predicted using (Q)SAR and the toxicity determined experimentally for the endpoint under investigation (Equation 11). Baseline toxicity is the lowest toxicity a chemical can exhibit, therefore a narcotic chemical would be expected to have a low TR, as long as the baseline toxicity prediction was accurate.

Equation 11 $$TR = \frac{LC/EC_{50,baseline}}{LC/EC_{50,experimental}}$$

where:

$LC/EC_{50,baseline}$:	baseline (non-polar narcotic) toxicity of a compound estimated using (Q)SAR (mol/L)
$LC/EC_{50,experimental}$:	toxicity of the compound determined experimentally (mol/L).

The approach proposed by Escher et al. (2006) allows estimation of the ecotoxicological range of a metabolite. The minimum or baseline ecotoxicity of a metabolite ($LC/EC_{50,baseline}$) is estimated using a suitable (Q)SAR, whilst the maximum (or specific) ecotoxicity ($LC/EC_{50,specific}$) is estimated by manipulating the baseline metabolite ecotoxicity with the TR of the parent compound (TR_{parent}) (Equation 12).

Equation 12 $$log\left(\frac{1}{LC/EC_{50,specific}}\right) = log\left(\frac{1}{LC/EC_{50,baseline}}\right) + log\, TR_{parent}$$

The estimation of $LC/EC_{50,specific}$ provides a worst-case estimate for metabolite ecotoxicity, that is, the metabolite has the same potency/mode of action as the parent PPP. Therefore the majority of predictions generated by this technique may overestimate the potency of a metabolite because there is an assumption that the metabolite(s) have the same potency as their parent PPP.

10.1.3.8. Other methods (from OECD (Q)SAR Application Toolbox)

OASIS acute toxicity mode of action profiler was developed by Professor A Zlatarov at the Laboratory of Mathematical Chemistry, University, Bourgas, Bulgaria. It is based on a broader set of structural alerts gathered primarily from the fathead minnow toxicity testing and defined by Russom et al. (1997).

The Verhaar classification (Verhaar et al., 1992) was developed utilising acute toxicity data collection for guppies and fathead minnows. This scheme based on structural alerts delineated chemicals into one of five classes. The Verhaar classes include (1) class 1 or 'inert' chemicals, which are non-polar narcosis or baseline toxicity; (2) class 2 or 'less inert' chemicals, which are the polar narcotics; (3) class 3 or 'reactivity' chemicals, which are typically non-selectively, covalently reactive with protein moieties; (4) class 4 or 'specifically-acting' chemicals, which specific reactivity with receptors; or (5) class 5 or 'unclassified' chemicals.

10.1.3.9. Read-across

Read-across for metabolites and impurities should include consideration of:

- molecular structure of the metabolite/impurity (active part intact or included?);
- the occurrence of metabolites in existing tests with the a.s. or major metabolites;
- general knowledge on the relationship between the toxicity of metabolites and their parent substances;
- available knowledge on related compounds.

10.1.4. Comparison of (Q)SAR model outputs

The accuracy of DEMETRA in predicting the acute toxicity of 135 PPPs to a standard aquatic test organism (*Daphnia magna*) was tested and compared with the performance of ECOSAR and TOPKAT. DEMETRA was found to provide more accurate predictions than ECOSAR and TOPKAT, which were not designed specifically for PPPs (Porcelli et al., 2008). The study indicated that ECOSAR (55 %) and TOPKAT (40 %) gave more false-negatives than DEMETRA (20 %). It should be noted that more PPP-specific classes have been added to ECOSAR since 2007.

Sinclair (2009) statistically compared measured and estimated acute toxicity data (n = 92) for metabolites of PPP on *Daphnia* applying DEMETRA, ECOSAR, TOPCAT and Escher et al. expert system. Results indicated that the simple expert system overall performed best. This is surprising as this approach is based on a relatively simple concept and indicates that transformation product toxicity is substantially linked to that of its parent PPP, or at least for transformation products within the evaluation dataset used in this case (Sinclair and Boxall, 2003). This comparison also indicated that DEMETRA performed better than TOPKAT and ECOSAR. Again, please note that more PPP classes have been added to ECOSAR after this comparison.

It is noted that the comparisons described above are hampered by the fact that the comparison of estimates and experimental toxicity was only based on acute toxicity data derived for *Daphnia*. Further validation should include other groups of aquatic organism.

Sinclair (2009) suggested that the simplest way of combining different estimation approaches would be to generate a conservative estimate of transformation product ecotoxicity, that is, estimating ecotoxicity using all approaches and then selecting the most potent prediction (see Figure 11a). Combining approaches in this manner would provide a conservative estimation of ecotoxicity. More sophisticated methods to aggregate model predictions may be considered in order to better handle outliers and increase predictive ability, for example, by calculating the geometric mean of predictive estimates or by developing rule-based methodologies. However, the likelihood of underestimating hazard may increase as can be seen from the calculation of geometric mean toxicity of five different

models in Figure 11b. Rule-based aggregation of predictions would require further investigation into (1) quantifying the predictive domain of each suitable approach, (2) rationalising the identity of outliers for each approach and (3) identifying which chemical types/categories are most appropriate for each approach. Developing such an approach would require a large transformation product dataset that extensively covers a range of taxa, physicochemical properties, transformation product chemical classes and parent pesticidal chemical classes. Such tasks are outside the scope of this GD. For the reasons above, the PPR Panel proposes to use the more conservative endpoint from different model estimates.

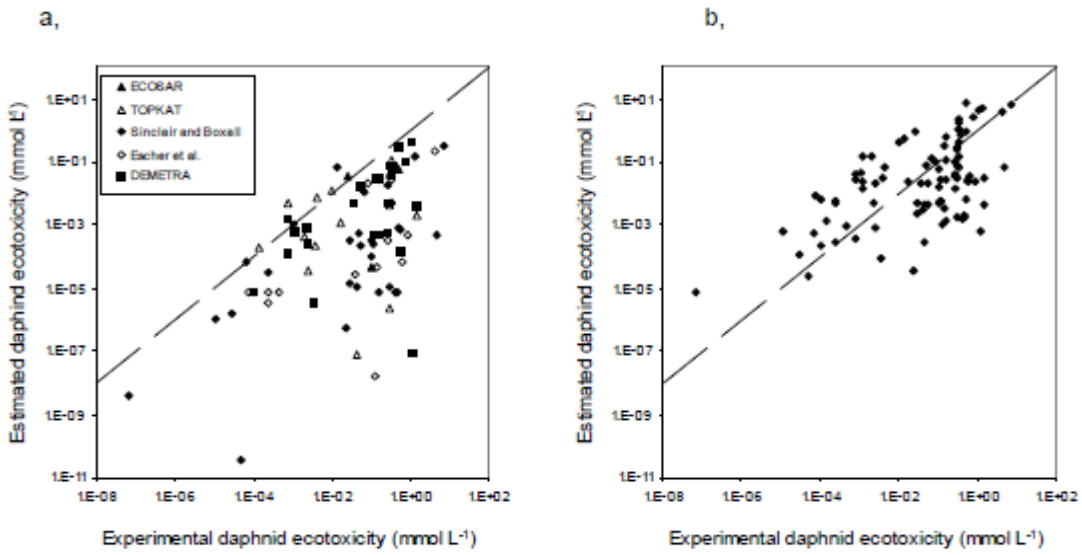


Figure 11: A comparison of daphnid acute ecotoxicity data for 92 transformation products; (a) the most potent estimates and (b), the geometric means provided by the five evaluated approaches (dashed line x=y) (from Sinclair, 2009)

10.1.5. Use of non-testing data in PPP risk assessment

10.1.5.1. General recommendations

The major concern in using non-testing data in environmental RA of PPPs is related to the danger of underestimating the real toxicity or hazard of a given substance.

No single model/expert system can be recommended for a substance, as the applicability of specific models depends on the adequacy in relation to this specific substance. However, estimates generated by the different approaches will vary, and it may therefore increase the likelihood that if a compound does exhibit a potent ecotoxicity for whatever reason, this will be picked up by at least one approach. It is further proposed to apply the model also on relevant analogues with experimental values in order to get a reasonability check for the model prediction.

Only suitable models (e.g. covering the right domain) with a high predictive reliability should be used. This should, among others, be reflected in the level of statistical significance required for estimates from (Q)SAR models. Validation parameters should ideally indicate good fits (e.g. $Q^2 > 0.7$, CCC > 0.85)[45]. Estimates of toxicity should where possible[46] be assisted by confidence intervals around the prediction. In case the standard deviation exceeds the predicted value itself, such values should not be accepted. Generally, the worst-case endpoint from several modelling approaches should be used.

[45] For further details consult ECHA (2008) guidance.
[46] Not all (Q)SAR models provide standard derivations for predictions.

Estimates should, however, be confirmed by using weight-of-evidence approaches where all available information is taken into account. This could include a combination of the different (Q)SAR model predictions combined with read-across and other available information like non-standard testing data and toxicodynamic/toxicokinetic information from mammals.

10.1.5.2. Modelling of impurities

Often ecotoxicological effect data are missing on impurities detected in new sources of a.s.. In case a non-testing approach is considered to provide effect data, the general recommendation above should be applied.

10.1.5.3. Modelling of metabolites

Whereas ecotoxicological effect data are normally present for the parent substance and major metabolites, such data are sometimes lacking for other metabolites. (Q)SAR may be used as a means to predict effect data to be used directly in assessment of risk and hazard, to prioritise metabolites of highest concern for further testing and/or to identify the likely most sensitive species for further testing (see section 10.2 on assessment of metabolites). The general recommendations above should be followed. The same level of strict validation criteria should be applied for long-term (Q)SAR ecotoxicity predictions as for acute (Q)SAR ecotoxicity predictions. It is noted that fewer valid (Q)SAR models are currently available for deriving longer term toxicity data.

10.1.6. Decision scheme for use of non-testing systems

1. Is the (Q)SAR model valid—i.e. is it relevant and reliable (following 5 OECD principles for assessing (Q)SAR models). For example, is prediction accurate enough (recommended assessment values Q^2, Concordance Correlation Coefficient (CCC), SD)?

   Yes: Go to 2
   No: (Q)SAR should not be used—consider other model.

2. Do the substance and model match—i.e. is the chemical of interest within the scope of the model? In order to do so, the following aspects should be considered:

   - Is the chemical in the applicability domain of the model, as described for the model?
   - Is the substance sufficiently similar to the compounds in the training set of the model (taking structural similarity but also and especially toxophores into account)?
   - Is the prediction for similar substances in the trainingset reliable (i.e. in line with the experimental data)?

   Yes: Go to 3
   No: (Q)SAR should not be used—consider other model.

3. Does model prediction take into account relevant substance properties (e.g. for aquatic toxicity consider water solubility, log K_{ow}, degradability and volatility)?

   Yes: Go to 4
   No: (Q)SAR should not be used—consider other model.

4. Are reliable estimations available from more than one (Q)SAR model?

   Yes: use lowest predicted (Q)SAR endpoint in RA or as qualifier for testing if confirmed by weight of evidence approach
   No: Single value could be used as qualifier for testing if clearly confirmed by weight of evidence approach.

10.2. Metabolites and degradation products

10.2.1. Introduction

Active substances in PPPs may be transformed in the environment by either abiotic or biotic processes. In Regulation (EC) No 1107/2009 a metabolite is defined as 'any metabolite or a degradation product of an active substance, safener or synergist, formed either in organisms or in the environment. A metabolite is deemed relevant if there is a reason to assume that it has intrinsic properties comparable to the parent substance in terms of its biological target activity [presence of toxophore], or that it poses a higher or comparable risk to organisms than the parent substance or that it has certain toxicological properties that are considered unacceptable. Such a metabolite is relevant for the overall approval decision or for the definition of risk mitigation measures'.

10.2.2. Definition of the residue for risk assessment

In the revised data requirement for a.s. (Commission Regulation (EU) 283/2013) it is stated under part A point 7.4.1 'Definition of the residue for risk assessment' that 'the residue definition relevant for risk assessment for each compartment shall be defined to include all components (active substance, metabolites, breakdown and reaction products) that were identified in accordance with the criteria referred to in this section. The chemical composition of residues occurring in soil, groundwater, surface water (freshwater, estuarine and marine), sediment and air, resulting from use, or proposed use, of a plant protection product containing the active substance, shall be taken into account'.

The criteria for identification are given in the section on fate and behaviour in the data requirements for the a.s. for the degradation in soil (Commission Regulation (EU) 283/2013 p. 7.1.1.) and for the degradation in surface water (Commission Regulation (EU) 283/2013 point 7.2.2.2. and 7.2.2.3) and cited below:

- 'identify the individual components present which at any time account for more than 10 % of the amount of active substance added, including, if possible, non-extractable residues;
- identify, if possible, the individual components which in at least two sequential measurements, account for more than 5 % of the amount of active substance added;
- identify, if possible, the individual components (> 5 %) for which at the end of the study the maximum of formation is not yet reached;
- identify or characterise, if possible, other individual components present'.

In addition to these requirements, the requirement for the degradation in surface water (Commission Regulation (EU) 283/2013 point 7.2.2.2., 7.2.2.3 and 7.2.2.4) indicates that the study shall 'permit, where relevant, the sediment residue of concern and to which non-target species are or may be exposed, to be defined'.

For metabolites that are formed at more than 10 % or between 5 and 10 % at two or more occasions or at more than 5 % at the end of the study, RA is needed. In general, RA for metabolites formed below 5 % or below 10 % (observed at a single occasion) is not considered necessary. However, if there is reason to believe that a metabolite formed at < 5 % has intrinsic properties comparable to the parent substance in terms of its biological target activity, or that it has certain structural properties indicating high reactivity (i.e. mutagenicity) or endocrine disrupting properties or that it has unacceptable toxicological properties, then that metabolite may be ecotoxicologically relevant and a RA is needed.

The following derogation (Commission Regulation (EU) 283/2013 point 7.1.1.2.2) from the above requirements applies for metabolites identified in the soil compartment:

'If, during adequate field studies metabolites, degradation and reaction products which are present in laboratory studies are below LOQ, which shall not exceed an equivalent of 5 % (molar basis) of the nominal concentration of a.s. applied, then in principle no additional information on the fate and behaviour of these compounds shall be provided. A scientifically valid justification for any discrepancy between laboratory and field appearance of metabolites shall be provided'.

Since no further information on fate and behaviour is necessary under these circumstances, it follows that no further information on ecotoxicity would be necessary.

In the revised data requirements (Commission Regulation (EU) No 283/2013) the lysimeter study is mentioned (point 7.1.4.2) as an experimental outdoor study in the framework of a tiered leaching assessment scheme. It is stated that the lysimeter studies shall be performed, where necessary, to provide information on the mobility in soil, the potential for leaching to ground water and the potential distribution in soil. It is not mentioned that the study should be used to identify metabolites relevant for ecotoxicological RA. It is required that the degradation and adsorption of metabolites that leach at a concentration level above 0.1 µg/L in the lysimeter is determined (Commission Regulation (EU) No 283/2013 point 7.1.1.2.1 and 7.1.2.1). This implies that metabolites only formed in a lysimeter are relevant for groundwater assessment but not for aquatic RA (see also point 9.2.4 of Commission Regulation (EU) No 284/2013 for PPPs). A leachate concentration of 0.1 µg/L in 100 mm percolating water corresponds with 0.0001 kg/ha, so 0.01 % of a dose of 1 kg/ha. So, triggering an aquatic RA for a metabolite based on exceedance of a lysimeter percolate concentration of 0.1 µg/L would be much more strict than triggering such an RA on the basis of exceeding 5–10 % of the dose of the parent in the soil, surface water (considering mineralisation, hydrolysis or photolysis) and water-sediment studies mentioned before. Moreover, such a lysimeter study is a higher tier study for the leaching assessment and thus not available on a standard basis. It seems therefore inconsistent to trigger an aquatic RA of a metabolite on the basis of the fact that it exceeds 0.1 µg/L in lysimeter percolate (of course unless there are ecotoxicological reasons to believe that this metabolite may cause a problem as described before).

If the metabolite is CO_2 or an inorganic compound that is not a heavy metal—or, it is an organic compound of aliphatic structure, with a chain length of four or less, which consists only of C, H, N or O atoms and which has no 'alerting structures' such as epoxide, nitrosamine, nitrile or other functional groups of known toxicological concern—then no further studies are required and the metabolite is considered to be not ecotoxicologically relevant and therefore of low risk to the environment.

All metabolites that, according to the criteria and definitions described above, are included in the 'Definition of the residue for RA' are hereafter called potentially relevant metabolites. For these, estimation of exposure (PEC) is necessary for each relevant compartment (see section 6.2.2), as well as information on ecotoxicity.

10.2.3. Identification of toxophore

Substances that have a specific mode of action, like PPPs, contain a structural feature or moiety that gives the toxic property. This structural feature is referred to as the toxophore, or toxophoric moiety. The substance causes toxicity through the interaction of its toxophore with a biomolecular site (e.g. receptor). Substances that are structurally similar could contain the same toxophore (or may yield a common toxophore upon metabolism) and may therefore have a common toxic effect.

For the assessment of the metabolite the applicant has to provide a reasoned case as to if the molecule contains a toxophore or if it has been lost following transformation. Toxophores for each of the major classes of PPP have been identified by looking for sub-structural similarities within a pesticidal class by Sinclair et al. (2009), which can be used to support argumentation. A number of ways have been identified to define the domain of applicability, which may be used to decide if toxophores are present or not (Nikolova and Jaworska, 2003; Dimitrov et al., 2005; Jaworska et al., 2005; Netzeva et al., 2005). In case it cannot be clearly shown that the toxophore is not present in the molecule, it should be

assumed that the toxophore remains and that the molecule has a specific mode of action (see assessment scheme 10.2.4).

10.2.4. Risk assessment scheme for metabolites

The decision scheme has been developed in order to facilitate the selection of the most appropriate and pragmatic assessment route for metabolites. However, possible endocrine disruption properties should be addressed separately (see section 3.3.6).

Sinclair and Boxall (2003) investigated the toxicity of metabolites in relation to the parent compound of several PPPs (60 a.s. and 485 transformation products) and demonstrated that the majority (70 %) of transformation products have either a similar toxicity to the parent compound or are less toxic. However, a significant proportion (30 %) were more toxic than their parent compound and 4.2 % of transformation products were more than an order of magnitude more toxic. Over 90 % of the observed increases in toxicity of the metabolite could be explained by the presence of a toxophore (see section 10.2.5), differences in accumulation (i.e. hydrophobicity) or differences in mode of action (for example active components of pro-PPPs or highly reactive metabolites). Furthermore, the investigation showed that transformation products that are more hydrophobic than their parent compounds and do not have pesticidal activity are unlikely to be more toxic than their parents to sensitive species that have a receptor site relevant to the parent mode of action. Hence, the PPR Panel has developed an assessment scheme, where, in the first step, metabolites for which it is clearly shown that the toxophore is lost can be assessed using approximation of toxicity (see section 10.2.7) while testing is required for metabolites with a remaining toxophore (see section 10.2.6).

1. Is the exposure to the metabolite in the toxicity test with the a.s. measured and adequate for assessing the potential effect of the metabolite (see section 10.2.5)?

   Yes: Go to 2
   No: Go to 3

2. Perform the RA assuming all the effect observed in the test with the a.s. can be attributed to the metabolite (see section 10.2.4, 10.2.5). Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Go to 3

3. Is it clear that the toxophore has been lost from the molecule (see section 10.2.3 and 10.2.7)?

   Yes: Go to 6
   No or unclear: Go to 4

4. Identify the species or taxonomic group[47] determining the lowest tier 1 $RAC_{sw;ac}$ for the a.s. Is the acute metabolite $L(E)C_{50}$> 10 times the a.s. $L(E)C_{50}$ (on a molar basis) (see section 10.2.6)?

   Yes: Go to 6
   No: Go to 5

5. Identify the species or taxonomic group[47] determining the lowest tier 1 $RAC_{sw;ch}$ of the a.s. Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Consider higher tier refinement

[47] Consider testing with tier 1 sediment organisms if metabolite is distributed in sediment (e.g. *Chironomus* for insecticides and *Lumbriculus* for fungicides).

6. Assume that the acute and chronic[48] toxicity of the metabolite is equal to the toxicity of the a.s. for all first tier taxonomic groups (see section 10.2.7). Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Go to 7 (for taxonomic groups where high risk is identified)

7. Are reliable and adequate non-testing predictions of toxicity (see section 10.2.8 and 10.1) available for all first tier taxonomic groups (fish, plants and invertebrates) for which risks were identified in step 6? Are $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$ using these predictions?

   Yes: Low risk
   No: Go to 8

8. Determine the acute and chronic[47,49] toxicity for those taxonomic groups for which risks were identified in step 6 and/or 7, and where a valid non-testing prediction of toxicity is not available or for which a risk was identified using predicted toxicity. Is $RAC_{sw;ac} > PEC_{sw}$ and $RAC_{sw;ch} > PEC_{sw}$?

   Yes: Low risk
   No: Consider higher tier refinement

10.2.5. Alternative information replacing experimental studies

The principles for assessing metabolites should, in essence, be the same as those for a.s.. However, in contrast to the a.s., data requirements for metabolites do not always have to be addressed by experimental studies. Applicants are invited to address the open questions by any other available information in support of a scientific and rational assessment. Examples of such information are shortly described below.

If chemical analyses confirm that the metabolite was present in the test system originally designed for testing of the a.s., organisms could be considered to have been exposed to the metabolites. The risk may then be addressed by information from the study with the a.s. assuming that all the observed effects in the test can be attributed to the metabolite when determining the RAC for the metabolite. However, this extrapolation is only valid if it is shown that the organisms were exposed to a realistic or worst-case exposure profile of the metabolite (e.g. compared with FOCUS profile or profile observed in the water sediment studies). For this extrapolation to be valid it is also important that the time period after the measured metabolite concentration was of sufficient length for observation of effects. In general, it will therefore only be possible to use the concentrations of the metabolite measured early in the test when establishing the RAC. Another possibility could be to prolong the test in order to lengthen the observation phase from effect occurring due to exposure to the metabolite.

If a metabolite is, for example, formed rapidly via hydrolysis, the toxicity of the metabolite may similarly have been assessed as part of the standard toxicity studies (this should be supported by analytical measurements) and addressed as above. However, if a toxicity test is performed at $pH \geq 7$ and if other metabolites are formed at pH 5 the toxicity of these metabolites needs to be addressed separately in order to cover the risk in more acid waters. Therefore the data from a hydrolysis study should be used to decide to what extent degradation and toxicity depend on the pH-value of the test medium.

In toxicity studies with intensive lighting (e.g. algae and *Lemna* tests), it is likely that metabolites which are formed as a result of photolysis are present in an amount which is relevant for field

[48] If chronic risk assessment is triggered by fate properties of the metabolite.
[49] For algae and aquatic plants, $RAC_{sw;ch}$ is based on EC_{50} values derived from chronic tests and no seperate $RAC_{sw;ac}$ is derived.

conditions and additional toxicity testing with metabolites detected in the photolysis study might not be warranted. This is particularly the case when static studies have been used. These conclusions should be supported by analytical measurements and the risk resulting from the metabolite can be addressed as above.

10.2.6. Metabolites structurally similar to the active substance and with remaining toxophore

It is likely that metabolites that are structurally similar to the a.s. (i.e. the toxophore remain) are most toxic to the same taxonomic group that was shown to be most sensitive to the a.s. For such compounds, testing can, in a first step, be limited to the taxonomic group that was identified to result in the lowest tier 1 $RAC_{sw,ac}$ and $RAC_{sw,ch}$ for the a.s. If, however, testing shows that this taxonomic group is not sensitive (i.e. the acute end point is greater than a factor 10 higher as compared with the parent, on a molar basis[50]) then it should be assumed that the molecule does not contain a toxophore. Hence, for the further assessment of this metabolite, it should be assumed that the most sensitive taxonomic group is unknown and the risk to all taxonomic groups should be addressed (see section 10.2.8).

If it is unclear if the toxophore remains and the most sensitive group is not known, then the RA needs to address all taxonomic groups.

10.2.7. Metabolites with no toxophore

If it is clear that the toxophore has been lost from the metabolite, in most cases, metabolites are less toxic to the target organisms than the a.s.. As a pragmatic and conservative approach for metabolites without the toxophore, the estimates of exposure could be compared with the RAC_{parent} based on the most sensitive endpoint of the a.s. in the relevant compartment. In general, only if this trigger is failed does the toxicity need to be further addressed (see section 10.2.8).

10.2.8. Non-testing predictions of metabolite toxicity

For metabolites which have lost the toxophore, the acute and long-term hazard and risk can be addressed using non-testing predictions of toxicity (see further in section 10.1). In principle, non-testing methods for predicting toxicity could also be used for specifically acting chemicals, that is, metabolites with a toxophore. However, based on the assessment by Sinclair and Boxall (2003) (see section 10.2.3) it is the view of the PPR Panel that only metabolites without toxophores should be included for non-testing estimates and is therefore expected to be of practical use only for a limited number of metabolites. If the trigger is failed using predicted toxicity then testing is required (see below).

10.2.9. Toxicity testing with metabolites

For metabolites with a possibly remaining toxophore, testing should, in the first step, be conducted with the species or taxonomic group[51], providing the lowest tier 1 $RAC_{sw:ac/ch}$ of the a.s. (see assessment scheme 10.2.4). For metabolites (without a toxophore) which require experimental studies (see assessment scheme in section 10.2.3), acute toxicity tests with *Daphnia*, rainbow trout and an algae should be conducted. In general, the same testing scheme as for a.s. (see Table 5.1) is required and hence only if the metabolite proves to be of similar toxicity as the a.s. should additional species also be tested.

[50] The statement to check whether the LC_{50} of the metabolite is greater than 10 times the LC of the a.s. on a molar basis means:

$$LC_{50_{met}} > 10\frac{M_{met}}{M_{ai}}LC_{50_{ai}}$$

where $LC50_{met}$ and $LC50_{ai}$ are mass concentrations (mg/L) of metabolite and a.s. at 50 % mortality and M_{met} and M_{ai} are the molar masses (g/mol) of the metabolite and a.s..

[51] Consider testing with tier 1 sediment organisms if metabolite is distributed in sediment (e.g. *Chironomus* for insecticides and *Lumbriculus* for fungicides).

In principle, for metabolites found in the sediment of a water-sediment study, the same triggers for testing should be applied to metabolites as for the a.s. (see section 7.2.5.1). That is, when accumulation of a substance in aquatic sediment is indicated or predicted from environmental fate studies, the impact on a sediment-dwelling organism should be assessed. Clearly the potential to exclude testing on the basis of toxicity will depend on the data available for the metabolite. The applicant should therefore make a case as to whether sediment testing can be waived based on what is known about the fate properties and toxicity profile of the metabolite. For example, if RAs with *Daphnia* indicate that the potential risks are low (taking into account the exposure situation in the sediment), then no further testing should, in general, be required. As a first screening step for metabolites partitioning to the sediment, a formula based on equilibrium partitioning theory as outlined in the TGD part II (EC, 2003) section 5.5.3, can be used to indicate if actual testing is needed. Only if a risk is indicated using this formula is actual testing with sediment organisms required. This will be further addressed in a PPR Panel opinion on sediment RA currently under development.

In order to decide whether chronic assessment is necessary, the intended uses and the fate and behaviour of the metabolite should be taken into account. In general chronic/long-term assessment are required for metabolites where exposure of surface water is likely and the metabolite is deemed to be stable in water, as defined in the data requirements, that is, there is less than 90 % loss of the original substance over 24 hours via hydrolysis under relevant pH conditions (Commission Regulation (EU) No 283/2013 and 284/2013). However, as hydrolysis studies are rarely available for metabolites, the 90 % loss trigger can be applied on data from other abiotic/biotic degradation studies.

For metabolites where chronic testing is necessary, the choice of taxonomic group(s) to be studied for chronic testing should take account of any acute toxicity data on the metabolite. Where information on the acute sensitivity of fish and invertebrates for a particular metabolite is available, chronic testing should only be required on the more sensitive group (i.e. that are a factor of 10 more sensitive). If *Daphnia* is suspected to be insensitive, based on the mode of action of the a.s. (e.g. it is an insect growth regulator or a neonicotinoid) then it is necessary to conduct a chronic study using the chironomid, *Chironomus riparius*, with the metabolite.

For unstable a.s. (i.e. there is more than 90 % loss of the original substance over 24 hours via hydrolysis), it may be more appropriate to conduct chronic studies on the stable metabolite instead of the parent compound. For unstable a.s., where chronic toxicity data for the parent compound are not available and a metabolite exceeds the persistence criteria (i.e. there is less than 90 % loss of the original substance over 24 hours via hydrolysis), chronic toxicity data should be submitted for this metabolite regardless of its acute toxicity.

The endocrine disrupting properties of metabolites should also be addressed; however, until common criteria are developed and agreed by the Commission, it is difficult to give specific guidance on how to assess endocrine disrupting compounds (EDC) in relation to PPPs (see section 3.3.6). Therefore, further guidance on how to assess EDCs might be given later when the work of the Commission is finalised. Nevertheless, based on structural properties of the metabolites and also based on information on related compounds indicating that the metabolite may exhibit endocrine disrupting properties, chronic/long-term tests with fish should always be required with this metabolite.

The BCF should be determined as for a.s. if the metabolite is stable (i.e. there is less than 90 % loss of the original substance over 24 hours via hydrolysis) and has a log P_{ow}> 3. In the first instance, QSARs could be used to predict the potential BCF. If appropriate information on the bioconcentration/bioaccumulation potential of the metabolite is available from parent BCF data, or other animal metabolism studies, this can be taken into account.

10.2.10. Risk assessment for metabolites

In principle, the RA process for metabolites will be similar to that for a.s., albeit recognising that RA cases will not always require specific study data for certain metabolites. If preliminary RAs indicate

potential concerns then, as for parent molecules, risk refinement is possible either by refining effect concentrations or by refinement of the exposure concentration.

If higher tier studies have been conducted with the a.s., or a relevant formulation, these studies may also have assessed the risk from the metabolites. It is advised that if a higher tier study, for example, a mesocosm study, is being carried out, then appropriate analysis should be conducted so that an assessment of both the exposure and effects of any metabolites can be made.

10.2.11. Definition of the residue for monitoring

Considering the results of toxicological and ecotoxicological testing, the residue for monitoring is defined to include only those components from the definition of the residue for RA which were classified as relevant in those tests (Commission Regulation (EU) No 283/2013 point 7.4.2). The PPR Panel proposes to define as relevant for monitoring only those metabolites for which a risk has been identified, that is, metabolites that pose a risk that warrants risk mitigation measures, in accordance with current practice.

For relevant metabolites, an analytical method for the compartments of interest should be available following Annex I listing.

10.3. Combinations of active substances in formulations

10.3.1. Background

The Regulation (EC) No 1107/2009 requires in Article 29 that 'interaction between the active substance, safeners, synergists and co-formulants shall be taken into account' in the evaluation and authorisation. This explicitly refers to marketed PPP, which are, by origin, technical mixtures containing one to several a.s., plus, typically, several co-formulants. Furthermore, the standard data requirements for PPP (Commission Regulation (EU) No 284/2013) do request 'any information on potentially unacceptable effects of the plant protection product on the environment, on plants and plant products shall be included as well as known and expected cumulative and synergistic effects'. First steps into guidance for an adequate consideration of mixture toxicity in the RA of PPP have been undertaken in the recently updated GD on birds and mammals (EFSA, 2009c) and in the scientific opinion on the science behind the development of an RA of PPPs on bees (EFSA PPR Panel, 2012b). The current proposal follows a similar approach, however, it is adapted to the typical data availability and specific approaches in the aquatic RA. More extensive background information might be obtained from (Kortenkamp et al, 2009; Altenburger et al, 2012; German Federal Environment Agency, 2013). The proposal aims at avoiding inconsistencies among EU Member States' RA. The focus of this section is on the technical mixtures of several a.s. and their co-formulants undergoing an authorisation procedure, solely.

In view of (i) the data typically available for the RA of PPP and their a.s., (ii) recent scientific opinions on the implementation of mixture RA in chemicals regulation (SCHER, SCCS, SCENHIR, 2012) and (iii) elements already applied and/or proposals currently brought forward by regulatory authorities of several European Member States (Altenburger et al., 2012; German Federal Environment Agency, 2013), two options are considered most adequate for the assessment of hazards and risks of pesticide mixtures under Regulation (EC) No 1107/2009 that involve measured and calculated mixture toxicity. As the intention is to improve mixture RAs without increasing testing requirements, the use of mixture toxicity calculations should be considered whenever justified (a priori, no synergistic effects) and possible (e.g. mixture composition of a.s. is different in the formulation than expected in the environment or experimental testing is technically not feasible).

10.3.2. Measured mixture toxicity

The most accurate option is to measure the mixture toxicity, that is, to perform experimental testing on the formulated product. According to the data requirements, tests on acute effects of the formulation should always be conducted for the most sensitive organism(s) (see section 7.5). In the case of

formulations with more than one a.s., this is justified as (i) the most sensitive organism to the given mixture of a.s. may not be easy to determine and (ii) co-formulants may contribute to the formulation toxicity. Furthermore, it is required to conduct chronic studies for formulations where the formulation is more acutely toxic than the a.s. by a factor of 10. To address this issue, we recommend the comparison of the measured acute ECx_{PPP} (endpoint for the formulation/PPP derived from experimental testing) and the acute $ECx_{mix\text{-}CA}$ calculated by CA (see sections 10.3.3 and 10.3.4). However, before requiring chronic testing, the possible contribution of co-formulants to the observed deviation between measured and calculated mixture toxicity should be carefully considered. If the toxicity of co-formulants is known, this can be included in the calculated $ECx_{mix\text{-}CA}$, and the comparison with the result from formulation testing (ECx_{PPP}) can be repeated.

10.3.3. Calculated mixture toxicity

In a recent review for the European Commission (Kortenkamp et al., 2009), the use of the concentration addition model (CA model) was proposed as the concept of mixture toxicity that is most relevant for hazard characterisation; it has also been extensively discussed by the European scientific committees (SCHER, SCCS, SCENIHR, 2012). The CA model is based on the following equation, for deriving a predicted EC_x or NOEC value for a mixture of (active) substances with known toxicity (EC_{xmix}-CA or NOECmix-CA), assuming concentration additivity:

Equation 13:
$$ECx_{mix-CA} = \left(\sum_{i=1}^{n} \frac{p_i}{ECx_i} \right)^{-1}$$

where:

- n: number of mixture components
- i: index from 1…n mixture components
- p_i: the i^{th} component as a relative fraction of the mixture composition (note: $\Sigma\, p_i$ must be 1)
- ECx_i: concentration of component i provoking x % effect (pragmatically, $NOEC_i$ may be inserted, too).

Two reasons make the use of the CA model concept attractive for regulatory purpose.

- it is generally more conservative than the concept of independent action (IA). Nevertheless, the magnitude of the differences at low levels of exposure between the two models is usually small and hence, the outcome will not be overly conservative;

- it can make use of existing data such as a NOEC, EC_{10} or EC_{50}s by applying the concept of TUs. The concept of TU has been recently reviewed by the three non-food committees of the EC (SCHER, SCCS, SCENIHR, 2012) which defined TU as 'the ratio between the concentration (i.e. ci) of a mixture component and its toxicological acute (e.g. EC_{50}) or chronic (e.g. long-term NOEC) endpoint'. In addition, the TU of a mixture has been defined as the sum of TU of each individual chemical of that mixture:

Equation 14:
$$\sum_{i=1}^{n} TU_i = \sum_{i=1}^{n} \frac{c_i}{ECx_i}$$

The committees also noted that the TU approach should refer to specific endpoints and to defined taxonomic groups of organisms (e.g. algae, daphnids and fish for the freshwater ecosystem) but not to the ecosystem as a whole (SCHER, SCCS, SCENIHR, 2012).

10.3.4. Counter-checking calculated and measured mixture toxicity

In order to determine if the a.s. may act more (i.e. synergistically) or less (i.e. antagonistically) than expected by CA, a comparison of the calculated $ECx_{mix\text{-}CA}$ for the mixture composition of a.s. in the formulation versus measured ECx_{PPP} endpoints is informative. This comparison may also indicate that relevant toxicity contributions of co-formulants not included in the calculation do occur, which might be included in a refined calculation (if the respective single-substance toxicity data are available). The deviation between calculated and measured mixture toxicity is—in line with (Belden et al., 2007a)—termed model deviation ratio (MDR):

Equation 15:
$$MDR = \frac{ECx_{mix-CA}\,(\text{calculated mixture toxicity})}{ECx_{PPP}\,(\text{measured mixture toxicity})}$$

- The observed and calculated mixture toxicities are considered in agreement if the MDR is between 0.2 and 5. This convention is in line with a proposal currently brought forward for the authorisation of biocidal products under the auspices of ECHA. In such a case, make use of the measured mixture toxicity (ECx_{ppp}) in the RA (at least, if mixture compositions in the study and at PEC_{mix} are compatible, see section 10.3.6).

- More-than additive (i.e. synergistic) mixture toxicity is indicated if the MDR is > 5. If synergistic effects cannot be excluded, the assessment should preferably be based on measurements (as synergistic interactions are not predictable by CA nor by other concepts such as IA). Regarding chronic testing requirements (see section 8.3.2) however, if the MDR for acute mixture toxicity falls between 5 and 10, a chronic study may not need to be performed. In those cases, a RA based on calculated mixture toxicity assuming CA might be considered if the relevant ETR trigger value is adapted in order to account for the observed synergism in the acute formulation study (at least should the ETR trigger be divided by the MDR). If experimental testing of the mixture is no option for certain species and endpoints (e.g. for technical reasons) but synergism cannot be excluded, the RA may be performed by adjusting the ETR as described above.

- Less-than additive (i.e. antagonistic) mixture toxicity is indicated if the MDR is below 0.2. If no plausible toxicological explanation for this apparent antagonism can be provided (e.g. special feature of the formulation type), the RA should be based on the calculated mixture toxicity.

A careful interpretation of the MDR is mandatory, especially if not all components that potentially contribute to the observed mixture toxicity (e.g. co-formulants) have been considered in the CA calculation. Furthermore, the outcome might be impacted by the heterogeneity of the toxicity data of individual mixture components used for a CA calculation and/or in comparison with the measured mixture toxicity data (e.g. differences in species, exposure designs, etc.), as discussed by Coors and Frische (2011). Care should also be taken that the counter-checking of measured and calculated mixture toxicity refers to the same basis, that is, the relative proportion of mixture components must be consistent (e.g. to the sum of a.s. of a given PPP if co-formulants are not included in the CA calculation). Additionally, for fluid PPP, it might be necessary to take the density into account in order to obtain exact figures.

10.3.5. Defining the mixture to be assessed

The total exposure concentration of the mixture (PEC_{mix}) is, in the first step, calculated as the simple sum of the PEC_i values of the n (number of components) individual components (per default: a.s.) by:

Equation 16: $$PEC_{mix} = \sum_{i=1}^{n} PEC_i$$

It should be carefully checked whether metabolites of (eco)toxicological relevance have to be included into the PEC_{mix} or not. For an initial screening approach, it is assumed that the $PEC_{sw;max}$ of all a.s. present in the formulation will occur at the same moment and are not separated in time (i.e. worst-case PEC_{mix}). In a subsequent step, more detailed consideration of the predicted exposure patterns in time can be undertaken to identify a more 'realistic worst-case' PEC_{mix} decisive for a refined mixture RA (see section below).

10.3.6. Risk assessment based on measured mixture toxicity

For a mixture RA based on measured mixture toxicity, the ETR is calculated by division of the PEC_{mix} divided by the measured mixture toxicity (ECx_{PPP}):

Equation 17: $$ETR_{PPP} = \frac{PEC_{mix}}{ECx_{PPP}}$$

It is, however, necessary to check whether the mixture composition in the formulation giving the measured toxicity (ECx_{PPP}) is similar to the mixture composition at the PEC_{mix} in terms of the relative proportions of the individual a.s.. As a direct comparison is not informative, as such, the comparison is done based on calculated mixture toxicity (assuming CA) for both mixture compositions, that is, a calculation of $ECx_{mix\text{-}CA}$ for the mixture composition of the a.s. at the PEC_{mix} and comparison with the respective estimate calculated for the formulation. The relative proportion of a.s. is considered sufficiently similar if the outcome of these calculations deviates less than 20 %. Hence, if ECx_{PPP} (proportion of a.s. as contained in PPP) divided by $ECx_{mix\text{-}CA}$ (proportion of a.s. at PEC_{mix}) yields a value between 0.8 and 1.2, a direct comparison of PEC_{mix} with the ECx_{PPP} is feasible. If the mixture composition differs more profoundly, the measured data cannot be used directly for calculating ETR; however, they might be used to justify the use of the calculated approach to perform the mixture RA (e.g. indicate the absence of more than additive effects).

10.3.7. Simplified approaches for mixture risk assessment

If no synergistic effects are indicated and the ETR values of the individual a.s. (ETRi) contained in the formulation are below the relevant trigger value, the mixture RA can follow a simplified approach: if all ETRi ≤ ETR trigger/n (n= number of a.s.) the mixture also fulfils the authorisation criteria and the procedure can be stopped. Care must be taken that the PECi values considered in ETRi are identical to those defined relevant for the mixture RA (i.e. PECmix = sum of PECi).

Furthermore, if the toxicity of the mixture is largely explained by the toxicity of a single a.s., a sufficient protection level might be achieved by simply basing the RA on the toxicity data for that single 'driver'. Hence, where CA provides a reliable estimate of the toxicity of the given mixture (ECx_{PPP}) and the largest part of the sum of toxic units (i.e. ≥ 90 %) calculated for the measured mixture toxicity (ECx_{PPP}) by Equation 14 comes from a single a.s., it can be concluded that this component drives the overall mixture toxicity (although formulation toxicity might be required, see section 7.5).

10.3.8. Risk assessment based on calculated mixture toxicity

For a mixture RA based on calculated mixture toxicity, the ETR is calculated by dividing PEC_{mix} by the calculated mixture toxicity assuming CA ($ECx_{mix\text{-}CA}$), which is the approach already followed by the regulatory authorities of several EU Member States (Altenburger et al., 2012). This assessment has to be carried out for each endpoint and exposure scenario separately unless it is evident that one specific endpoint/exposure scenario combination clearly drives the risk. The scheme as shown is

focusing on effect concentrations (ECx) but may equally well be applied to NOEC data if these are considered as low effect concentrations. $ETR_{mix\text{-}CA}$ is defined by:

Equation 18: $$ETR_{mix-CA} = \frac{PEC_{mix}}{ECx_{mix-CA}}$$

To do so, the transformed CA equation (Equation 13) is plugged into the ETR_{mix} definition (Equation 18), with the total exposure concentration of the mixture (PEC_{mix}, see section 10.3.5) and the relative proportions of the individual mixture components (p*i*) defined as

Equation 19: $$p_i = \frac{PEC_i}{PEC_{mix}}$$

which finally gives:

Equation 20: $$ETR_{mix-CA} = \frac{PEC_{mix}}{ECx_{mix-CA}} = \frac{\sum_{i=1}^{n} PEC_i}{\frac{1}{\sum_{i=1}^{n} \frac{p_i}{ECx_i}}} = \sum_{i=1}^{n} ETR_i, \quad if \ p_i = \frac{PEC_i}{PEC_{mix}} = \frac{PEC_i}{\sum_{i=1}^{n} PEC_i}$$

If the standard acceptability criteria based on first-tier data (i.e. standard laboratory data) and worst case PEC_{mix} are met ($ETR_{mix} \leq$ ETR trigger value), the risk from a CA action of the mixture is considered acceptably low. If the relevant trigger value is not met, further available refinement options regarding both exposure and effect assessment might be checked on a case-by-case basis. In case the endpoints to be used for the mixture RA refer to the same taxonomic group but are associated with different AFs (e.g. single species test, Geomean or SSD), the calculation of the mixture toxicity could not be based on the ETR_{mix} but instead on the regulatory acceptable concentration of the individual a.s. (RAC*i*) using the following formula yielding a risk quotient for the mixture (RQ_{mix}):

Equation 21: $$RQ_{mix} = \sum_{i=1}^{n} \frac{PEC_i}{RAC_i}$$

If $RQ_{mix} < 1$, the risk is considered acceptable.

The use of this formula hence allows the use of refined endpoints from, for example, SSD in the mixture toxicity calculations. However, the concepts of CA have been developed and evaluated for single species mixture toxicity assessments (tier 1 standard laboratory data). It is unclear whether and under what conditions they are also applicable to endpoints derived from experimental ecosystem studies (e.g. micro-/mesocosms). Furthermore, the concept should only refer to data from the same relevant taxonomic group (i.e. fish, invertebrates, algae and aquatic plants), but not to the ecosystem as a whole. Therefore, it could be argued that in accordance with SCHER, SCCS, SCENIHR (2012) it is recommended not to use data from micro-/mesocosms (e.g. ETO-RAC) until further knowledge and proper guidance is available. However, the pragmatic use of ETO-RAC from experimental ecosystem studies is currently used by several Member States for calculated mixture toxicity assessment.

10.3.9. Independent action for mixture toxicity calculation

Details on the possibility of performing a mixture assessment using IA are not presented here, as current knowledge does not support a full implementation of the IA option in this guidance for routine mixture RA. However, there is no principal scientific reservation against the application of IA under Regulation (EC) No 1107/2009 if the necessary information and input data are available. Further

information on the IA approach might be obtained from Kortenkamp et al. (2009). In order to check whether the IA option might theoretically lead to the conclusion of an acceptable risk, an examination of the maximal possible difference between corresponding mixture toxicity predictions by means of CA and IA according to Junghans et al. (2006) can be conducted for the defined mixture (PEC_{mix}). The examination is based on the toxic unit approach (Equation 14) applied to a specific toxicity endpoint and the mixture as defined by the PEC_{mix}:

$$ETR_{mix-CA} \leq trigger \times \frac{\sum_{i=1}^{n} \frac{PEC_i}{ECx_i}}{\max\left\{\frac{PEC_i}{ECx_i}\right\}}$$

Equation 22:

Where this criterion is fulfilled, the assumption of IA may finally lead to the conclusion of an acceptable risk and thus the IA option might be worth being considered in more detailed.

10.3.10. Possibilities to refine the worst-case PEC_{mix}

As a first approach, it is assumed that the $PEC_{mix;max}$ of all a.s. present in the formulation will occur at the same moment and are not separated in time. In case the trigger value is not met, the predicted exposure patterns can be taken into account in refinement steps, that is, by considering the differing profiles of PECi of the individual a.s. when entering the water bodies (see examples in Table 38 and

Table **39**) Such refinement options are especially relevant as the exposure patterns of a.s. can vary according to the entry route, that is, spraydrift, run-off/erosion and drainage. In such cases, the changes in relative proportions of the individual a.s. in the mixture composition over time are considered in the calculation because the maximum concentrations of each a.s. may not occur simultaneously (e.g. drainage). The assessment needs to be repeated for all relevant FOCUS scenarios.

Table 38: Example for a tier 1 calculation using highest peaks (worst-case PEC_{mix}) for a mixture of two compounds (all concentrations in µg/l) for acute first tier risk assessment of fish or invertebrates)

Concentration highest peak compound A	0.12
Concentration highest peak compound B	0.23
Concentration highest peak A + B: PEC_{mix}	0.35
Toxicity compound A	10.0
Toxicity compound B	8.00
Toxicity mixture	8.59
ETR mixture	**0.041 * (unacceptable risk)**

*Values in bold > trigger value of 0.01; additional risk assessment should be considered.

Table 39: Example of refinement of $PEC_{mix;\ max}$ for a mixture of two compounds (all concentrations in µg/l). Values printed in bold are above the trigger value of 0.01 and additional risk assessment should be considered (example for acute first tier risk assessment of fish or invertebrates)

Days	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Concentration compound A	0.00	0.12	0.09	0.08	0.07	0.06	0.05
Concentration compound B	0.23	0.12	0.06	0.03	0.01	0.00	0.00
Concentration compounds A & B	0.23	0.24	0.15	0.11	0.08	0.06	0.05
Toxicity compound A	10.00	10.00	10.00	10.00	10.00	10.00	10.00

Days	1	2	3	4	5	6	7
Toxicity compound B	8.00	8.00	8.00	8.00	8.00	8.00	8.00
Toxicity mixture	8,00	9,00	9,20	9,45	9,75	10	10
ETR mixture	**0,029***	**0,027***	**0,016***	**0,011***	0,008	0,006	0.005

*Values in bold > trigger value of 0.01; risk unacceptable.

10.3.11. Decision scheme for mixture toxicity risk assessment

Note, this mixture RA scheme has to be carried out for each endpoint and exposure scenario separately unless it is evident that one specific endpoint/exposure scenario combination clearly drives the risk. The scheme as shown is focusing on effect concentrations (ECx), but may equally well be applied to NOEC data if these are pragmatically considered as low effect concentrations.

1. Are measured toxicity data (ECx) available for the given endpoint (typically chronic data available only for a.s.)?

   Only for the a.s. ($ECx_{a.s.}$): Go to 7
   For both formulation (ECx_{PPP}) and a.s. ($ECx_{a.s.}$): Go to 2

2. Check the plausibility of the measured formulation toxicity (ECx_{PPP}) against the calculated mixture toxicity $ECx_{mix\text{-}CA}$ (assuming CA, Equation 13) for exactly the mixture composition of the a.s. in the formulation (ECx_{PPP}) by means of the model deviation ratio (MDR = $ECx_{mix\text{-}CA}/ECx_{PPP}$).

   If MDR = 0.2–5 (CA approximately holds for the mixture): Go to 3
   If MDR > 5 (mixture more toxic than CA): Go to 10
   If MDR < 0.2 (mixture less toxic than CA): Go to 9

3. Check whether the mixture composition in the formulation study giving the measured mixture toxicity (ECx_{PPP}) in terms of the relative proportions of the individual a.s. is similar to the mixture composition at the PEC_{mix}[52]. As a direct comparison on the basis of the relative proportions of the a.s. at the ECx_{PPP} with the relative proportion at the PEC_{mix} is not informative as such, the comparison is done based on calculated mixture toxicity (assuming CA) for both mixture compositions. Therefore, calculate $ECx_{mix\text{-}CA}$ (see Equation 13) for the mixture composition of the a.s. at the PEC_{mix} and compare with the estimate calculated for the formulation (as already done in step 2 above).

   If $ECx_{mix\text{-}CA}$ (a.s. in PPP)/$ECx_{mix\text{-}CA}$ (a.s. in PEC_{mix}) = 0.8–1.2 (mixture similar): Go to 4
   If not (mixture not similar): Go to 5

4. Conduct a mixture RA based on measured mixture toxicity, with the exposure-toxicity ratio (ETR_{mix}) being defined as the PEC_{mix} divided by the measured ECx_{PPP} and compare the outcome with the acceptability criterion (trigger value) decisive for the specific endpoint/exposure scenario combination.

   If ETR_{mix} < trigger: Low risk
   If ETR_{mix} > trigger: low risk not demonstrated/check refinement options

[52] Define the mixture to be assessed in terms of the relative proportions (p_i) of the individual mixture components (i) at the PEC_{mix} with p_i being defined as the PEC of the individual components (PEC_i) divided by PECmix. For an initial screening consider per default the $PEC_{sw\ max}$ of the individual active substances contained in the formulation (i.e. PEC_{mix} equals sum of PEC_i). Additionally check whether metabolites of ecotoxicological relevance have to be included into the PEC_{mix} or not).

5. Check whether one mixture component clearly drives the toxicity if considering the measured mixture toxicity (ECx_{PPP}), that is, does the largest part of the sum of toxic units (Equation 14) calculated for the formulation (≥ 90 %) comes from a single a.s. (TU_i)[53]?

   Yes (single 'driver' of mixture toxicity identified): Go to 6
   No: Go to 8

6. Conduct a RA based on single-substance toxicity data ($ECx_{a.s.}$) for the identified 'driver' of mixture toxicity, with the exposure-toxicity ratio ($ETR_{a.s.}$) being defined as the $PEC_{a.s.}$ divided by the measured $ECx_{a.s.}$ and compare the outcome with the acceptability criterion (trigger value) decisive for the specific endpoint/exposure scenario combination.

   If $ETR_{a.s.}$ < trigger: Low risk
   If $TER_{a.s.}$ > trigger: low risk not demonstrated/Check single-substance refinement options

7. Is there evidence that synergistic interactions between mixture components might occur (e.g. based on toxicological knowledge from literature or from counter-checking measured and calculated mixture toxicity in other species) which cannot be ruled out for the given species with sufficient certainty?

   Yes (mixture toxicity calculation not feasible): Measured mixture toxicity data required for RA (if becoming available: Go to 2)
   No (mixture toxicity calculation feasible): Go to 8

8. Conduct a mixture RA based on calculated mixture toxicity according to 10.3.8:

$$ETR_{mix-CA} = \frac{PEC_{mix}}{ECx_{mix-CA}}$$

   If $ETR_{mix\text{-}CA}$ < trigger: Low risk
   If $ETR_{mix\text{-}CA}$ > trigger: Low risk not demonstrated, check single-substance refinement options

   If the endpoints to be used for the RA refer to the same taxonomic group but are associated with different AFs (e.g. single species test, Geomean or SSD), the calculation of the mixture risk is assessed by:

$$RQ_{mix} = \sum_{i=1}^{n} \frac{PECi}{RACi}$$

   If RQ < 1: Low risk
   If RQ > 1: Low risk not demonstrated/check exposure refinement options (see 10.3.10)

9. Carefully recheck the apparent antagonism as observed in the measured mixture toxicity data (ECx_{PPP}) regarding potential impacts of the default assumption of CA and/or heterogeneous input data used for the CA calculation. Does the apparent antagonism remain and no toxicologically plausible explanation is available (e.g. special feature of the formulation type)?

   Yes (measured mixture toxicity not plausible): Go to 8

[53] with TU*i* being defined as the concentration of the i^{th} a.s. at the ECx_{PPP} (re-caculated to the sum of a.s.) divided by the respective single-substance toxicity ($ECx_{a.s.}$).

No (measured mixture toxicity plausible): Go to 3

10. Carefully recheck the apparent synergism as observed in the measured mixture toxicity data (ECx_{PPP}) regarding potential impacts of heterogeneous input data (a.s.) and of co-formulants ignored in the CA calculation. Does the apparent synergism remain?

    Yes: Go to 3, if measured data are not available (see section 7.5.2), or if the assessment in point 3 indicates that the mixtures are not similar, **go to 8** (use modified ETR trigger values, see 10.3.4)
    No: Go to 3

11. Other issues

11.1. Test batches/impurities

Differences in the chemical composition of the test batches might alter the ecotoxicological profile of the technical material. In Regulation (EC) No 1107/2009, impurities are defined as any component other than the pure a.s. and/or variant which is present in the technical material (including components originating from the manufacturing process or from degradation during storage) (Art. 3 (33)). In the EC GD SANCO/10597/2003, rev. 9 of 17 June 2011 for assessing the equivalence of technical material, an impurity is considered significant when it occurs, due to process variability, in quantities ≥ 1 g/kg in the a.s. as manufactured, based on dry weight; an impurity is considered relevant when it is of toxicological and/or ecotoxicological or environmental concern compared with the a.s., even if present in technical material at < 1 g/kg.

Information on the composition of the test batches used in the ecotoxicological tests should be available. When the composition of the batches is comparable with the specification of technical material (SANCO/10597/2003), consisting of the a.s. and its associated impurities, no further assessment is requested. However, when the composition of the batches is different, in terms of different amount of the same impurities or different impurities or different amount of the a.s., then the ecotoxicological representativeness of the related endpoints should be addressed. To determine whether the batches used in the ecotoxicity studies are equivalent to those which will be approved, it is recommended that the assessment methodology provided in the latest version of the EC GD SANCO/10597/2003 is followed. It is recommended that a table with the ecotoxicity tests, the batches analysed and the observed endpoints is included to make the comparison of the toxicity of different batches easier.

11.2. Testing poorly soluble and other difficult test substances

Detailed guidance on how to deal with poorly soluble substances as well as other substances that are difficult in aquatic toxicity testing (e.g. volatile or adsorbing substances) can be found in the 'OECD Guidance document on aquatic toxicity testing of difficult substances and mixtures' (OECD, 2000). For poorly soluble substances, limit concentrations lower than 100 mg/l may also be acceptable (see OECD, 2000). Precipitation of the substance in the test medium should be avoided because data generated under these circumstances are usually highly variable. It is generally not sufficient to test the maximum water solubility of the substance because this is usually determined in studies with pure water under sterile conditions. Attempts should be made to reach the maximum solubility level expected under the test conditions, where necessary using a solubiliser, solvent vehicle or dispersing agent, though care should be taken in respect of alterations in bioavailability or confounding effects, particularly in chronic studies (Hutchinson et al., 2006). For some compounds, the solubility in pure water is likely to be higher than in standard test media. If, on the basis of these results a potential risk is identified (from the appropriate ETR), further testing may be necessary. Studies on the formulated product might also be an appropriate way to deal with poorly soluble compounds, especially if no effects occur at the solubility limit.

11.3. Promising mechanistic effect models

Mechanistic ecological models have been applied to ecotoxicological questions for over 25 years now (e.g. O'Neill et al., 1982, Kooijman and Metz, 1984), but over the years their acceptance in regulatory environmental risk assessment (ERA) has been rather limited. Current RA is based mainly on statistical (e.g. LOEC, LC_{50}, SSD) and physical models (e.g. *Daphnia magna*, microcosm) which do not explicitly rely on systematic understanding on the system of interest. However, more recently, the importance of mechanistic modelling has increased, especially in the ERA of PPPs under the European Regulation (EC) No 1107/2009. Mechanistic modelling was mentioned as a valuable higher tier tool in the SETAC workshops AMPERE on mesocosm tests (2007) and AMRAP on macrophyte testing (Maltby et al., 2010), the ELINK workshop on complex exposure scenarios (Brock et al., 2010a) and finally the LEMTOX workshop discussing pros and cons of population models (Thorbek et al., 2010). Another SETAC workshop (EU Workshop on how to use ecological effect models to link ecotoxicological tests to protection goals (MODELINK)) was held in 2012/2013, though proceedings for this were not available before the finalisation of this GD. Additionally mechanistic modelling is explicitly mentioned in the EFSA opinion on protection goals (EFSA PPR Panel, 2010a).

Despite the potential power of mechanistic effect modelling to answer important questions within ERA and its long history in science, its use within ERA is not well tested nowadays and no general guidance is available. In the near future, the PPR Panel will elaborate scientific opinions on good modelling practice (EFSA-Q-2011–00989) and more specifically on modelling within the aquatic RA (EFSA-Q-2012–00960). Since there is a lack of experience and guidance for these approaches in RA, the use of mechanistic modelling within the authorisation of PPPs has to be evaluated carefully case-by-case until special guidance becomes available.

It is expected that mechanistic effect models at all levels of biological organisation will be used to support the RA of PPPs in the future. On the individual level, for example, TK/TD models simulate survival (Jager et al., 2011) as well as sublethal effects over time, based on uptake and elimination of the toxicant (toxicokinetics) and damage and repair processes (toxicodynamics) within the organism (Jager et al., 2006). Population models aim at extrapolating lethal and sublethal effects from the individual to the population level (Forbes and Calow, 2002; Barnthouse, 2004; Van den Brink et al., 2007; Preuss et al., 2010). Ecosystem models (Hommen et al., 1993; Traas et al., 2004; Park et al., 2008; De Laender et al., 2013) allow the risk characterisation within ecosystems, integrating biotic interactions and thereby indirect effects as reviewed for experimental studies (Relyea and Hoverman, 2006). Population and ecosystem modelling may serve as useful tools through which, for example, protection goal fulfilment of lower tier RA can be explored; however, due to the current state of development of these models, detailed recommendations for their use in RA cannot be given at this stage.

11.4. Reduction of (vertebrate) testing

Directive 2010/63/EU on the protection of animals used for scientific purposes, describes that 'when choosing methods, the principles of replacement, reduction and refinement should be implemented through a strict hierarchy of the requirement to use alternative methods'.

Regulation (EC) No 1107/2009 clearly requires 'the use of non-animal test methods and other RA strategies should be promoted. Animal testing for the purposes of this Regulation should be minimised and tests on vertebrates should be undertaken as a last resort'. Therefore, aquatic RA alternatives to experimental testing are specifically recommended for fish. The OECD recently revised the Fish Toxicity Testing Framework also in this respect (OECD, 2012b).

11.4.1. Use of limit tests

The data requirements (Commission Regulation (EU) No 283/2013) generally state for aquatic organisms that a limit test at 100 mg substance/L may be performed when the results of a range finding test indicate that no effects are to be expected.

Specifically with respect to minimising vertebrate/fish testing, a threshold approach to acute fish testing should be considered. An acute fish limit test should be conducted at 100 mg substance/L or at an appropriate concentration selected from aquatic endpoints following consideration of the threshold exposure (see also OECD guideline on acute fish testing 203). For the limit concentration to be tested for poorly soluble substances, see also section 9.2. When mortality is detected in the fish limit test, an acute fish dose–response toxicity study shall be required to determine an LC_{50} for use in RA (see also chapter 7).

A workshop was held in December 2010 in the UK to investigate the possibilities of using the threshold approach for acute fish testing in PPP RA (Creton et al., in preparation). The threshold approach was considered useful for assessing a.s.. Furthermore, possibilities for reducing the number of fish in bioconcentration studies were recently presented (Creton et al., 2013).

11.4.2. Use of non-testing methods

Non-testing methods such as (Q)SAR and read-across can be applied to fill certain data gaps if no test data are available. This particularly applies to testing for metabolites or impurities where appropriate non-testing methods can be recommended in certain cases (see section 10.1 and 10.2). In particular, it is recommended to invoke consensus modelling strategies through combining predictions from methods of different types such as read-across, (Q)SAR and structural alert models. In this way, consensus outcomes will indicate an increased level of confidence, while dissent outcomes may indicate the need for more data before proceeding with the RA.

11.5. Differences in risk assessment procedures between Regulation (EC) No 1107/2009 and the Water Framework Directive (WFD)

11.5.1. Introduction

In Europe, different legislations (Directives, Regulations) have been developed with different methodologies to assess the aquatic risks of PPPs. In particular, these differences are apparent when comparing the authorisation criteria for the compartment water according to the PPP Regulation (EC) No 1107/2009 and the water quality standards according to the WFD (2000/60/EC). These criteria and standards are a reflection not only of differences in the use of data on environmental fate and ecotoxicology of PPPs, but also of different policy decisions about the acceptance of risks in relation to formulated protection goals.

The WFD aims to maintain and improve the aquatic environment in EU Member States so that a ‘good ecological status’ and a ‘good chemical status’ is achieved. For a good status, the WFD requires that environmental quality standards (EQSs) are met. These EQSs are one of the instruments to evaluate water quality and serve as a benchmark to decide whether or not specific measures are required. A distinction is made in the annual average (AA)-EQS and the maximum acceptable concentration (MAC)-EQS. The AA-EQS aims to protect aquatic ecosystems and organisms from effects owing to long-term exposures. The MAC-EQS aims to protect aquatic ecosystems and organisms from short-term concentration peaks. The methodology for EQS derivation is described in the ‘Technical Guidance for Deriving Environmental Quality Standards under the Water Framework Directive’ (EC, 2011b).

During the approval of a.s. at EU level, the relevant data to derive EQSs need to be compiled. The data requirements (Commission Regulation (EU) 283/2013) state that ‘all of the aquatic toxicity data shall be used when developing a proposal for environmental quality standards (Annual Average EQS, AA-EQS; Maximum Acceptable Concentration EQS, MAC-EQS). The methodology for derivation of these endpoints is outlined in the “Technical Guidance for Deriving Environmental Quality Standards” for the Water Framework Directive 2000/60/EC’. It should be noted that in proposing an EQS, the endpoints from some standard and higher tier studies are interpreted and used in a different way from that conventionally used in RAs.

11.5.2. Overview of the main differences in risk assessment procedures between plant protection product regulation and the Water Framework Directive

11.5.2.1. Chemical context

Under the umbrella of the WFD, EQSs are derived for all toxic chemicals (e.g. metals and all types of organic pollutants) that are identified as problematic in (one of the) European river basins. In contrast, Regulation (EC) No 1107/2009 exclusively deals with the RA of PPPs used (or intended for use) in the EU.

11.5.2.2. Protection goals

The protection goals underlying the WFD refer to human and ecosystem health. Within the context of ecosystem health and the EQS setting it is assumed that (1) ecosystem sensitivity depends on the most sensitive species (population) and (2) protecting ecosystem structure protects community functioning. EQSs are derived on the basis of predicted no effect concentrations (PNECs) for all relevant populations of water organisms (comparable to the ETO). Although the generic protection goals of the WFD and PPP Regulation do not differ substantially, the SPGs of the Plant Protection Product Regulation do not exclude that under certain conditions short-term effects followed by recovery are acceptable (ERO), while EQS setting within the context of the WFD in principle is based on the ETO.

11.5.2.3. Geographical context

The aquatic RA procedure, according to the WFD, has its focus on usually larger water bodies within the context of river basins. The aquatic RA procedure under the umbrella of the PPP Regulation has its focus on edge-of-field surface waters in agricultural landscapes. In some EU Member States, a clear differentiation of non-WFD and WFD water bodies is implemented.

Exposure assessment

The RA, according to the WFD, follows a retrospective approach. The chemicals evaluated are already used (placed on the market) and form potential problems in one or more European water basins. Comparing chemical monitoring data, based on analysis of discrete chemical monitoring samples, with the AA-EQS and MAC-EQS is the means by which compliance is assessed. According to the 'Technical Guidance for Deriving Environmental Quality Standards under the Water Framework Directive', EQSs should be linked to an annual average concentration (AA-EQS) or the maximum of the measured concentrations (MAC-EQS).

The RA procedure according to the PPP Regulation follows a prospective approach. This approach allows assessment of the risks of a PPP before it is placed on the market. A common, cost-effective approach in the prospective exposure assessment is the use of harmonised exposure scenarios (FOCUS Surface Water Scenarios). These scenarios, in combination with models that estimate the emissions to and the fate and behaviour of PPPs in surface waters, intend to predict realistic worst-case exposure concentrations in edge-of-field surface waters. The $RAC_{sw;\ ac}$ is compared with the $PEC_{sw;max}$ and the $RAC_{sw;ch}$ is, in first instance, compared with the $PEC_{sw;max}$ and under certain conditions with the $PEC_{sw;twa}$. The time window for the $PEC_{sw;twa}$ is usually smaller than the duration of the standard toxicity test that triggered the risk in tier 1. The way of linking exposure to effects is substantially different in the WFD approach.

11.5.2.4. Effect assessment

In the effect assessment, the PPP Regulation follows a tiered approach while the EQS derivation, according to the WFD, follows a weight-of-evidence approach. The main differences in effect assessment approaches concern (1) the use of toxicity data for algae and macrophytes, (2) the use of additional toxicity data (e.g. to construct SSDs) and (3) the way micro-/mesocosm are used.

1. In the WFD approach, the EC_{50} values for algae and macrophytes are used in the acute effect assessment and the NOEC/EC_{10} values in the chronic effect assessment. In the tier 1 effect assessment of the PPP regulation, the EC_{50} values of algae and macrophytes are used (in both the acute and chronic assessment).

2. If, besides the base set, additional toxicity data are available, it is possible to apply the SSD approach. To apply the SSD approach, the procedures developed for the PPP Regulation require at least five (fish) to eight (plants and invertebrate) toxicity data points for different taxa of the sensitive taxonomic group. The SSD approach developed to derive WFD EQSs requires at least 10 toxicity values for at least eight different taxonomic groups. So, in the first instance, the specific toxic mode of action of the PPP will not be considered when constructing the SSD for an EQS setting. The SSD procedure developed for RAC derivation (PPP Regulation) and WFD EQS derivation is based on calculating the HC_5 and by applying an AF. The height of the AF, however, differs for RAC derivation and EQS derivation. If the number of additional toxicity data is less than required for the SSD approach the WFD methodology will select the lowest toxicity value to derive the EQS by applying an AF. For RAC derivation (PPP Regulation) the Geomean approach may be used.

3. For EQS derivation the threshold levels for effects derived from appropriate micro-/mesocosm tests may be used by applying an appropriate AF. For RAC derivation, on the basis of appropriate micro-/mesocosm tests, both the 'ETO' and the 'ERO' may be followed. In RAC derivation, the predicted exposure profile for the edge-of-field surface water of concern plays a prominent role when interpreting results of micro-/mesocosm studies.

A detailed description of EQS derivation procedures can be found in the 'Technical Guidance for Deriving Environmental Quality Standards under the Water Framework Directive' (EC, 2011b). In Brock et al. (2011) a description is given on how the EQS derivation can be performed for PPPs, while also the differences between the WFD and PPR Regulation procedures are explained in greater detail.

12. Addressing uncertainties

12.1. Approaches for characterising uncertainty in higher tier assessments

Regulation (EC) No 1107/2009 lists under Annex II criteria for approval of a.s., safeners and synergists under 3.8 'Ecotoxicology', point 3.8.1 '…The assessment must take into account the severity of effects, **the uncertainty of the data**, and the number of organisms groups which the a.s., safener or synergist is expected to affect adversely by the intended use'. This implies that uncertainties in the data must be considered.

Regulation (EC) No 546/2011 states that no authorisation shall be granted unless it is 'clearly established' that no unacceptable impact occurs. The term 'clearly established' implies a requirement for some degree of certainty. First tier assessments use standardised scenarios and decision rules which are designed to provide an appropriate degree of certainty. Higher tier assessments are not standardised, and so the degree of certainty they provide has to be evaluated case-by-case. The need for RAs to include characterisation of uncertainty has also been emphasised at senior policy levels in the EU[54] (see also Sterling, 2010).

Methods for characterising uncertainty can be grouped into three main types:

- Qualitative methods: using words to describe the certainty of an outcome, or to describe how different the true outcome might be compared with an estimate.

[54] E.g. 'Even though it is not a subject that lends itself easily to quantification, I would urge you to take account of the risk manager's need to understand the level of uncertainty in your advice and to work towards a systematic approach to this problem.' (Madelin, 2004).

- Deterministic methods: generating deterministic quantitative estimates of impact for a range of possible scenarios. This shows the range of possible outcomes (e.g. a range of ETRs) and can be accompanied by qualitative descriptions of their relative probabilities (traditional 'worst-case' assessments are an example of this).
- Probabilistic methods: these give numeric estimates of the probabilities of different outcomes. These probabilities may be estimated statistically (e.g. when quantifying measurement or sampling uncertainty, or as outputs from probabilistic modelling). However, they may also be estimated subjectively, by expert judgement.

All uncertainties affecting an assessment should be considered, at least qualitatively. To reduce the risk of overlooking important uncertainties, it is recommended to systematically consider each part of the assessment (e.g. different lines of evidence, different inputs to calculations, etc.) and list all of the sources of uncertainty together with a description of the magnitude and direction of their potential influence on the expected level of impact. As well as evaluating each individual source of uncertainty, it is also essential to give an indication of their combined effect. It is recommended to use a tabular approach to facilitate and document this process, as illustrated in Table 40. This is based on an approach used in some PPR Panel opinions (EFSA, 2006b, 2007c, d, 2008), but adapted to increase clarity by introducing separate columns to describe uncertainties that act in different directions.

Research in social science has shown that there is a general tendency for experts to underestimate uncertainties (Morgan et al., 2009). It is therefore important that risk assessors are aware of the potential magnitude of common uncertainties in the assessment of risks to non-target organisms. For example, assessors should be aware of the potential magnitude of measurement uncertainties (e.g. abundance of daphnids in water samples from mesocosms and of the potential magnitude of sampling uncertainty associated with small and moderate sized datasets).

In some cases, a qualitative evaluation of uncertainties may be sufficient to establish clearly (i.e. with sufficient certainty) that unacceptable levels of impact will not occur, as is required by the 'unless' clause in Regulation (EC) No 546/2011. In other cases, a purely qualitative evaluation of uncertainty may not give a sufficiently clear picture of the range of possible outcomes. In such cases, one option is to obtain additional data to reduce uncertainty. This may usefully be targeted on the uncertainties that appeared largest in the qualitative evaluation. However, an alternative option is to refine the characterisation of the uncertainties progressively by evaluating some of them, first using deterministic methods and then, if necessary, probabilistic methods. This implies a tiered approach to the treatment of uncertainties, which starts by evaluating all uncertainties qualitatively and progresses either by reducing uncertainty (by obtaining additional data) or by refining the evaluation of selected uncertainties (either deterministically or probabilistically), until the point where it can be 'clearly established' whether an unacceptable impact will occur (as required by the 'unless' clause in Regulation (EC) No 546/2011).

Table 40: Tabular approach recommended for qualitative evaluation of uncertainties in refined assessments. The +/– symbols indicate whether each source of uncertainty has the potential to make the true risk higher (+) or lower (–) than the outcome of the refined assessment. The number of symbols provides a subjective relative evaluation of the magnitude of the effect (e.g. +++ indicates an uncertainty that could make the true risk much higher). If the effect could vary over a range, lower and upper evaluations are given (e.g. +/++). If possible, the user should indicate the meaning of different numbers of symbols (e.g. two symbols might be used to represent a factor of 5, and three symbols a factor of 10). See Appendix G for practical examples

Source of uncertainty	**Potential to make true risk lower**	**Explanation**	**Potential to make true risk higher**	**Explanation**
Concise description of first source of uncertainty	Degree of negative effect	Short narrative text explaining how this factor could make true		

Source of uncertainty	Potential to make true risk lower	Explanation	Potential to make true risk higher	Explanation
	(e.g. – – –)	risk lower		
Second source of uncertainty			Degree of positive effect (e.g. +++)	Short narrative text explaining how this factor could make true risk higher
Add extra rows as required for additional sources of uncertainty	–	Note: many uncertainties may act in both positive and negative directions	+	
Overall assessment	**Narrative text describing the assessor's subjective evaluation of the overall degree of uncertainty affecting the assessment outcome, taking account of all the uncertainties identified above and a conclusion concerning the direction the overall uncertainty leads to (true risk higher or lower). The overall assessment should be a balanced judgement and not simply a summation of the plus and minus symbols**			

It is unlikely that it will ever be practical—or necessary—to quantify all uncertainties, so every deterministic or probabilistic assessment should be accompanied by a qualitative evaluation of the unquantified uncertainties. Moreover, it should be remembered that deterministic and probabilistic methods often require assumptions (e.g. about distribution shapes) that are themselves uncertain, and these additional uncertainties should be included in the qualitative evaluation. Therefore, every refined assessment should contain at least a qualitative evaluation of uncertainties.

The overall magnitude of uncertainty associated with an assessment will often be very large. This should not be regarded as implying a failure of RA; on the contrary, it provides essential information for decision making (Madelin, 2004; Sterling, 2010).

It should be noted that for PPPs where several different types of refined assessment are used, the uncertainties affecting each one will be different. In such cases, it is recommended to evaluate the uncertainties affecting each approach separately. The contribution of the multiple assessment approaches (multiple lines of evidence) in reducing overall uncertainty can then be evaluated by weight-of-evidence in the final risk characterisation (see next section).

In summary, it is recommended that:

- Every refined RA should be accompanied by at least a qualitative evaluation of the uncertainties affecting it, using a systematic tabular approach. In assessments with multiple lines of evidence, the uncertainties affecting each line of evidence should be evaluated separately.
- In cases where qualitative evaluation of uncertainty is not sufficient to determine whether it is clearly established that no unacceptable impact will occur, the assessor may either (a) seek further data to reduce the uncertainty, or (b) refine the evaluation of the existing uncertainties using quantitative methods (which can be either deterministic or probabilistic).

12.2. Risk characterisation and weight-of-evidence assessment

Risk characterisation is the final step of RA. At this point, all relevant information or evidence that has been gathered is used to produce an overall characterisation or description of the risk, in a form that is suitable for decision making.

To be useful for decision making, the risk characterisation should focus on evaluating whether the relevant protection goals are satisfied for the PPP under assessment. Often, risk characterisation will

involve combining several different types of refined assessment, each providing a separate indication of the risk. This process of combining available 'lines of evidence' to form an integrated conclusion or risk characterisation is frequently referred to as 'weight-of-evidence' assessment (e.g. EC, 2002c; Hull and Swanson, 2006). This term reflects the principle that the contribution of each line of evidence should be considered in proportion to its weight. This weight-of-evidence assessment should also consider the presumptions of the tiered approach as adopted in this AGD (see chapter 4).

In the context of this document, a line of evidence might be the completed output of any of the refinement options as described in chapters 8 and 9.

The PPR Panel recommends a qualitative[55] approach to weight-of-evidence assessment, as follows:

- Consider all relevant lines of evidence, including the first tier assessment. Retention of the first tier assessment is appropriate in all cases, as it is relevant to consider whether it was borderline or failed by a large margin.
- Evaluate the uncertainties associated with each line of evidence. This should be done by applying the approaches described in the preceding section to each line of evidence separately. The characterisation of overall uncertainty for each line of evidence is then used in the weight-of-evidence assessment, as in principle the weight given to each line of evidence should be proportionate to its certainty.
- Form overall conclusions by using expert judgement to combine all lines of evidence, weighted according to their certainty, and give more weight to the most certain, but also take due account of the less certain. High certainty implies high weight. If one line of evidence implies a much narrower range for the risk than another line of evidence (i.e. higher certainty), then the true risk is most likely to fall inside the range of the former.
- Be sure to take full account of the uncertainties and to include a fair description of the range of possible outcomes in the final risk characterisation. Identify the outcome that is considered most likely, but do not give it more emphasis than is justified by the evidence.
- If different lines of evidence conflict (e.g. effects in mesocosm studies at concentrations lower than the tier 1 RAC), this should be considered a form of uncertainty. No line of evidence should be completely discounted unless it is wholly invalid or irrelevant. Instead, as stated above, each line of evidence should contribute to the overall conclusion in proportion to its certainty.
- If the overall characterisation of risk is expressed qualitatively, choose words very carefully to describe the outcome and its uncertainty as clearly as possible. For example the phrase 'on balance' is often used to focus on one of several possible outcomes, for example, 'on balance, it is concluded there will be no mortality'. This type of statement is not appropriate, because it fails to communicate the degree of certainty (e.g. 'on balance' could mean 51 % or 99 % certainty%)[56].
- A weight-of-evidence assessment is inevitably subjective. Different assessors may vary in their weighing of the evidence, especially when uncertainty is high. Therefore, it is essential to document the assessment in detail, including the outcome and uncertainty for each line of evidence considered, and explaining how they were combined to reach conclusions about the overall outcome and its uncertainty.

The PPR Panel recommends a systematic tabular approach to documenting the weight-of-evidence assessment, such as that illustrated in

Table **41**. The tabular format provides a concise yet clear summary of the lines of evidence considered and how they were combined. It also helps the reader to evaluate whether the assessment was balanced and aids consistency of approach between PPPs.

[55] Quantitative approaches could also be used to combine lines of evidence, but this requires each line of evidence to be expressed in the same units together with a quantitative measure of its certainty.
[56] Note that the standard of evidence required by the 'unless' clause is 'clearly established', which is much stronger than 'on balance'.

It should be noted that

Table **41** summarises the major types of uncertainty for each line of evidence, and not just the overall uncertainty. This is recommended because it helps the assessor to take account of some important strengths and weaknesses of different types of refined assessment (EFSA, 2009c).

The subjectivity of weight-of-evidence assessment can impede the formation of an independent view when this is based on the assessment of another person. Therefore, when a weight-of-evidence assessment is submitted by an applicant, it would be prudent for the regulatory authority to conduct their own weight-of-evidence assessment separately, compare their conclusion with that of the applicant and consider the reasons for any differences.

It is sometimes argued that characterising uncertainty is unhelpful in decision making. In fact, it is essential for risk assessors to characterise uncertainty, as is clear from Regulation 546/2011 ('clearly established') and from policy statements by the EC (Madelin, 2004). Furthermore, practical options exist for dealing with uncertainty in decision making. Two of the principal options are to request more data to reduce uncertainty or to request more refined evaluation or analysis of the existing uncertainty. A third option is to counter the uncertainty by applying risk mitigation options, so that the chance of adverse impacts is limited to an acceptable level.[57] However, choosing between options for dealing with uncertainty involves risk management considerations outside the scope of this document such as the acceptability of effects, the degree of certainty required and potentially other factors such as the cost and time required for further refinement, the need to respect legal deadlines for authorisations, and the consequences of risk mitigation or non-authorisation (e.g. reduced efficacy, reduced choice of pest control options in agriculture, risk of resistance, etc.).

In summary, the PPR Panel recommends that:

- Every refined RA should conclude with an overall characterisation of risk, in terms relevant for decision making. It is recommended to begin with the consideration of whether the evidence makes any effects unlikely that are in conflict with the SPG for the water organisms under evaluation. Where this is not satisfied, attention should turn to characterising the levels of unacceptable effects that may occur, and using this to evaluate whether there is a high certainty of no short-term and/or long-term repercussions on, for example, acute mortality of aquatic vertebrates and abundance and diversity of populations of aquatic vertebrates, invertebrates and primary producers (the actual protection goal).
- The overall characterisation of risk should be derived by a qualitative weight-of-evidence assessment considering all relevant lines of evidence and their uncertainties using a systematic tabular approach (e.g.
- Table **41**). If the overall characterisation is expressed qualitatively (in words) rather than quantitatively, great care should be taken to describe the outcome and its uncertainty as clearly as possible.
- The first tier assessment should always be included as one of the lines of evidence and should be given appropriate weight.

Table 41: Tabular approach recommended for qualitative weight-of-evidence assessment, summarising the conclusion and uncertainties for several lines of evidence and using them to develop an overall conclusion. See Appendix G for practical examples. The +/– symbols indicate whether each source of uncertainty has the potential to make the true risk higher (+) or lower (–) than the indicated outcome. The number of symbols provides a subjective relative evaluation of the magnitude of the

[57] 'In cases where both the potential risk and scientific uncertainties are high, the risk manager may conclude that a precautionary approach is appropriate.' (Madelin, 2004).

effect (e.g. – – – might indicate an uncertainty that could reduce risk by an amount equivalent to reducing an ETR by about a factor of 10). If the effect could vary over a range, lower and upper evaluations are given (e.g. –/++ or +/++)

	Lines of evidence *(add more columns if appropriate)*		
	First tier assessment *(should always be included)*	**Second line of evidence**	***Add one column for each line of evidence***
Main contributions to uncertainty			
Concise description of first major source of uncertainty	+ and – symbols (see legend)		
Second uncertainty			
Add one row for each major source of uncertainty			
Conclusions for individual lines of evidence	Insert overall assessment for each line of evidence		
Overall conclusion	**Insert overall conclusion giving appropriate weight to each line of evidence, taking account of their relative certainty (more uncertainty = less weight). The overall conclusion should be a balanced judgement and not simply a summation of the plus and minus symbols**		

12.3. Uncertainties in extrapolating to real field situations

To meet the requirements of a protective approach in RA, it is necessary to not only predict, but also retrospectively monitor the effects of pesticides in the field as done for every credible model going beyond a 'Gedankenexperiment' (theoretical exercise). Available information on unacceptable effects of pesticides in the field indicates that care has to be taken using the RA approach used until now, as effects have been identified. This includes:

- The EU/SETAC workshop EPiF identified that 'effects of pesticides were identified in several of the field studies' (Liess et al., 2005).
- In Australia, pesticide effects on invertebrates were identified in a stream following run-off (Muschal and Warne, 2003).
- In Germany, effects of pesticides on an invertebrate community were observed following run-off. Possible confounding factors as hydrodynamic stress were experimentally excluded (Liess and Schulz, 1999). Additionally, in several other streams, pesticide effects were identified (Liess and von der Ohe, 2005).
- In South Africa, Dabrowski et al. (2001) identified pesticide effects on a stream.
- A recent meta-study of Schäfer et al. (2012) identified pesticide effects in eight data sets comprising exposure data and effect data.

Possible reasons for the effects observed, and solutions to solve the problems identified, could be:

- Non-compliance with good agricultural practice and risk mitigation measures so that the exposure is higher than predicted.
- In the current exposure assessment, certain exposure routes might not be covered.
- Effects are not covered by the current prospective RA (first tier or higher tier).

- The possible shortcomings of the prospective RA procedure should, at least in part, be addressed by an appropriate implementation of the WFD and the Sustainable Pesticide Use Directive. If, in chemical monitoring programmes, certain pesticides (regularly) exceed the WFD water quality standards (AA-EQS; MAC) this should have consequences for the re-registration if it can be demonstrated that the problems are caused by use of a specific pesticide, or other appropriate measures should be taken (e.g. punishing the farmers that do not follow label instructions; implementing more strict risk mitigation measures). The results presented above may give important evidence showing that, in practice, the RA based on existing methodology is, in some cases, not protective enough for aquatic non-target communities in edge-of-field water bodies. If there is some concern about the safety of a product, competent authorities could request chemical as well as biological post-authorisation monitoring in edge-of-field water bodies. Results from this monitoring could be considered in relation to the authorisation of the substance. Note that WFD environmental quality standards are stricter than the RACs, since the WFD also aims to protect aquatic communities in larger surface waters and different procedures in the linking of exposure to effects are used. It may be a future option in the registration procedure to always require a chemical (and biological) monitoring programme in edge-of-field surface waters for a few years if a new substance is placed on the market. Appropriate monitoring programmes can, however, only be set up once the exposure assessment goals are defined.
- Environmental stress may alter effects of toxicants on populations and communities by a factor of more than 10. Examples include investigations by Foit et al. (2012), Knillmann et al. (2012a), Liess and Beketov (2011), Reynaldi et al. (2011), and Stampfli et al. (2011). In addition to the effect, recovery may also be influenced (Foit et al., 2012; Knillmann et al., 2012b).
- A mixture of the possibilities mentioned above.
- Surely another plausible explanation for the failure of RA to have predicted/precluded these effects, is that the effects reported were not a result of single pesticide exposure. This is in the realm of ecoepidemiology, and similar uncertainties apply with regard to isolating potential causative factors and eliminating confounding factors. Examples within the framework that certainly will underestimate the risk are tank mixtures, while multiple exposure (serial application of several PPP, multi-crop, multi-year) and long-term delayed effects of pesticides may be underestimated in the RA procedure described in this GD.

Establishing a firm link between the exposure concentration of a single pesticide and its effect in the field faces the problem that, mostly, several pesticides occur simultaneously in streams. However, in small agricultural streams (and other edge-of-field surface waters) one or two substances are, in most of the cases, strongly dominating the toxicity at the same time and/or in sequence. This has been shown in field experiments using realistic application rates of the total package of pesticides used in a wheat crop (Auber et al., 2011), a potato crop (Arts et al., 2006) and a tulip crop (Van Wijngaarden et al., 2004) and in field monitoring studies of agricultural areas in Australia (Muschal and Warne, 2003), Germany (Liess and Schulz, 1999; Liess and von der Ohe, 2005), France (Schäfer et al., 2007) and the USA (Belden et al., 2007b). Hence, observed effects are, in many cases, related to the effects of one or two a.s. (that may be characterised by repeated pulse exposure) in one year. This may often concern substances that differ in toxic mode of action (e.g. insecticide and fungicide) if effects are caused by more a.s..

Within this document, some sources of uncertainty are not considered. These include the uncertainty related to. for example, (i) multiple exposure to different pesticides; (ii) multi-year sequential pesticide exposure; (iii) combined effects between the PPP and environmental stressors as hydrodynamic stress (e.g. unfavourable temperature) (for i–iii see also EFSA PPR Panel, 2012b); and (iv) especially sensitive species such as stoneflies (Wogram and Liess, 2001; Ohe and Liess, 2004) that are generally not used in test systems as they are difficult to culture. The uncertainties mentioned here are relevant for all tiers in RA and need to be considered in addition to the AFs identified here. Future investigations and guidance needs to identify the magnitude of this uncertainty.

12.3.1. Conclusions

- Long-lasting effects of pesticides were identified in several field studies. Therefore, as mentioned above, uncertainty should be taken into account and if there is some concern about the safety of a product, competent authorities could request chemical and possibly biological post-authorisation monitoring in edge-of-field water bodies.

- As long as there is no guidance on how to include additional stressors into the RA, we need to reflect the related uncertainty. Hence it should be realised that, in the first tier, an AF of 100 for acute and 10 for chronic toxicity may not be protective in 100 % of the cases. Consequently this will also be the case for all effect tiers. Therefore, there is a need to validate/calibrate the RA scheme to the field situation. When there is a systematic deviation for substances with a specific mode of action, this should trigger a revision of the RA scheme.

- Care should be taken when extrapolating from tier 1 and higher tiers (such as mesocosms and surrogate reference tiers) to the field, which is the ultimate reference tier. This extrapolation will benefit from ecological modelling (when available in the future) including all relevant processes necessary for extrapolation (including environmental stress and biological interaction).

12.3.2. Research needed

- Field investigations need to exemplarily verify exposure and effect predictions (see also Artigas et al., 2012).

- The link between results obtained by tier 1 tests with the situation in the field need to be strengthened. This includes investigations on the degree to which effects from single species tests can be altered by the environmental context.

- The link between results obtained by mesocosms with the situation in the field need to be strengthened. This includes investigations on the degree to which effects from field and mesocosms can be altered by the environmental context.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

A new tiered RA scheme for aquatic organisms in edge-of-field surface water is presented that updated the former GD on Aquatic Ecotoxicology (EC, 2002a) taking into account the new Regulation (EC) No 1107/2009, new data requirements (Commission Regulations (EU) No 283/2013 and 284/2013) and relevant scientific publications and findings. Detailed scientific background for the proposed assessment can be found in chapters 3 to 12, whereas the short guidance which is intended for regular day-to-day use is presented in chapter 2.

RECOMMENDATIONS

- The AFs for tier 1 as given in the uniform principles (Regulation (EC) No 546/2011) of 100 on acute toxicity data and 10 for chronic toxicity data have not been sufficiently validated/calibrated for all types of PPPs and it is not fully clear whether all relevant uncertainties are covered in any case. The PPR Panel therefore recommends that a future scientific opinion elaborates the scientific justification for the tier 1 AFs.

- The current scheme has been elaborated and checked based on modes of actions of a.s. that have so far undergone the regulatory RA. For new a.s. with new modes of action, it needs always to be checked carefully whether the RA scheme addresses all potential risks, for example, determining the most appropriate test organisms by performing a battery of tests on a broad range of species.

- As outlined above, the GD was checked for some a.s. for which a larger data set is available. However, the PPR Panel would recommend the further checking of the GD on a reasonable number of a.s. to check for consistency once experience with the new GD has been gained.

- The current GD suggests several new methodologies for the evaluation of micro-/mesocosm studies in order to facilitate high-quality studies (e.g. introducing the use of the minimum detectable difference (MDD) and community-level effect analysis by trait based methods). Currently there is only limited experience available with these methods and after more experience on their use has been gained this should be re-evaluated.

- An evaluation of differences in overall sensitivity and recovery potential of sensitive populations between lentic (ditches and ponds) and lotic (streams) micro-/mesocosms is recommended.

- Since currently available toxicity data for aquatic macrophytes are characterised by a high variability in endpoints assessed, it is recommended to evaluate their ecological significance for effect assessment at the population and community level (aim: cost-effective and protective approach).

- Calibrations were performed between higher and lower tier data for invertebrates in this GD. The outcomes of this calibration were extrapolated to the cases with macrophytes and fish. A proper calibration between different tiers for macrophytes should be performed in the future.

- It could be shown that the acute toxicity data for rainbow trout can reasonably cover aquatic life stages of amphibians. However, for a suitable refinement in the RA, it needs to be investigated further whether there are differences in the SSDs of fish and amphibians.

- There is currently limited experience on the use of the Geomean approach and of SSDs on chronic data due to limitations in data availability. Some validation exercise is recommended to demonstrate whether the approaches are equally applicable to chronic data on the basis of similar endpoints.

- The RA is performed for single substances and the question on how protective the RA is regarding the actual use patterns in the field is arising. Consequences of simultaneous and sequential exposure to different PPPs as well as stress due to environmental stressors should be further investigated to better address issues of (i) the 'uniform principles' as laid down in Regulation (EC) No 546/2011 that requires that Member States base their authorisation decision on the 'proposed conditions for the use of the PPP' and (ii) the standard data requirements for PPP do request that 'any information on potentially unacceptable effects of the PPP on the environment, on plants and plant products shall be included as well as known and expected cumulative and synergistic effects'.

- The current scheme has been elaborated on on the basis of calibrations between higher and lower tier data for invertebrates. The higher tier data are from a set of micro-/mesocosm studies which represent the 'surrogate reference tier'. With the reference being the field itself, it is recommended to conduct further investigations in the field to strengthen the links between results obtained in the tiered approach and the situation in the field, that is, to perform a retrospective evaluation.

- The PPR Panel did not revise or evaluate the current exposure assessment but advises to critically evaluate and improve the surface water exposure assessment in the future.

- For RA based on calculated mixture toxicity, the concept of CA has been developed and evaluated for single species mixture toxicity assessments (tier 1 standard laboratory data), and it is unclear whether and under what conditions it is also applicable to endpoints derived from experimental ecosystem studies (e.g. ETO-RAC from micro-/mesocosms). This is an area for further research.

- It needs to be discussed with risk managers whether the potential risks of PPP exposure to aquatic microbes (bacteria, fungi) need to be addressed, and, if so, which tier 1 data requirement should be adopted.

REFERENCES

Abel PD, 1980. Toxicity of hexachlorcyclohexane (Lindane) to Gammarus pulex; mortality in relation to concentration and duration of exposure. Freshwater Biology, 10, 251–259.

Adriaanse PI and Beltman WHJ, 2009. Transient water flow in the TOXSWA model (FOCUS versions): concepts and mathematical description. Statutory Research Tasks Unit for Nature and the Environment, WOT report 101, Wageningen, The Netherlands.

Aldenberg T and Jaworska JS, 2000. Uncertainty of hazardous concentrations and fraction affeceted for normal species sensitivity distributions. Ecotoxicology and Environmental Safety, 46, 1–18.

Aldenberg T, Jaworska JS and Traas TP, 2002. Normal species sensitivity distributions in probabilistic ecological risk assessment. In: Species sensitivity distributions in risk assessment. Eds Posthuma L, Traas TP and Suter GW. CRC Press, Boca Raton, FL USA, 49–102.

Alonso Prados E and Novillo-Villajos A, 2010. Ecological characterization of permanent and ephemeral streams of a typical Mediterranean agricultural landscape (east and southeast of Iberian peninsula). In: Linking aquatic exposure and effects: risk assessment of pesticides. Eds Brock TCM, Alix A, Brown CD, Capri E, Gottesbüren BFF, Heimbach F, Lythgo CM, Schulz R and Streloke M. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 288–303.

Aldrich AP, 2009. Empfindlichkeit von Amphibien gegenüber Pflanzenschutzmitteln. AGRARForschung, 16, 466–471.

Altenburger R, Arrhenius A, Backhaus T, Coors A, Faust M and Zitzkat D, 2012. Ecotoxicological combined effects from chemical mixtures—Part 1: Relevance and adequate consideration in environmental risk assessment of plant protection products and biocides (Project No. (FKZ) 3709 65 404). Umweltbundesamt, Dessau-Rosslau, Germany).

Artigas J, Arts G, Babut M, Caracciolo AB, Charles S, Chaumot A, Combourieu B, Dahllöf I, Despréaux D, Ferrari B, Friberg N, Garric J, Geffard O, Gourlay-Francé C, Hein M, Hjorth M, Krauss M, De Lange HJ, Lahr J, Lehtonen KK, Lettieri T, Liess M, Lofts S, Mayer P, Morin S, Paschke A, Svendsen C, Usseglio-Polatera P, van den Brink N, Vindimian E and Williams R, 2012. Towards a renewed research agenda in ecotoxicology. Environmental Pollution, 160, 201–206.

Arts GHP, Buijse-Bogdan LL, Belgers JDM, Van Rhenen-Kersten CH, Van Wijngaarden RPA, Roessink I, Maund SJ, Van den Brink PJ and Brock TCM, 2006. Ecological impact in ditch mesocosms of simulated spray drift from a crop protection programme for potatoes. Integrated Environmental Assessment and Management, 2, 105–125.

Ashauer R, Boxall ABA and Brown CD, 2007. New ecotoxicological model to simulate survival of aquatic invertebrates after exposure to fluctuating and sequential pulses of pesticides. Environmental Science and Technology, 41, 1480–1486.

Auber A, Roucate M, Togola A and Caquet T, 2011. Structural and functional effects of conventional and low pesticide input crop-protection programs on benthic macroinvertebrates communities in outdoor pond mesocosms. Ecotoxicology, 20, 20142–2055.

Baird DJ and Van den Brink PJ, 2007. Using biological traits to predict species sensitivity to toxic substances. Ecotoxicology and Environmental Safety, 67, 296–301.

Bärlocher F, 1985. The role of fungi in the nutrition of stream invertebrates. Botanical Journal of the Linnean Society, 91, 83–94.

Barnthouse LW, 2004. Quantifying population recovery rates for ecological risk assessment. Environmental Toxicology and Chemistry, 23, 500–508.

BBA, 2000. Bekanntmachung über die Abdrifteckwerte, die bei der Prüfung und Zulassung von Pflanzenschutzmitteln herangezogen werden. (8. Mai 2000) in: Bundesanzeiger No.100, amtlicher Teil, vom 25. Mai 2000, S. 9879.

Becker RA, Janus ER, White RD, Kruszewski FH and Brackett RE, 2009. Good Laboratory Practices and Safety Assessments. Environmental Health Perspectives, 117, 482–483.

Beketov MA and Liess M, 2005. Acute contamination with Esfenvalerate and food limitation: Chronic effects on the mayfly *Cloeon dipterum*. Environmental Toxicology and Chemistry, 24, 1281–1286.

Beketov MA and Liess M, 2008. Acute and delayed effects of the neonicotinoid insecticide thiacloprid on seven freshwater arthropods. Environmental Toxicology and Chemistry, 27,461–470.

Belden JB, Gilliom RJ and Lydy MJ, 2007a. How well can we predict the toxicity of pesticide mixtures to aquatic life? Integrated Environmental Assessment and Management, 3, 364–372.

Belden JB, Gilliom RS, Martin JD and Lydy MJ, 2007b. Relative toxicity and occurence patterns of pesticide mixtures in streams draining agricultural watersheds dominated by maize and soybean production. Integrated Environmental Assessment and Management, 3, 90–100.

Belgers JDM, Aalderink GH, Arts GHP and Brock TCM, 2011. Can time-weighted average exposure concentrations be used to assess the risks of metsulfuron-methyl to *Myriophyllum spicatum*? Chemosphere, 85, 1017–1025.

Bergtold M and Dohmen GP, 2011. Biomass or growth rate endpoint for algae and aquatic plants: relevance for the aquatic risk assessment of herbicides. Integrated Environmental Assessment and Management, 7, 237–247.

Biggs J and Brown CD, 2010. Ecological characterization of water bodies in clay landscapes in the United Kingdom. In: Linking aquatic exposure and effects: risk assessment of pesticides. Eds Brock TCM, Alix A, Brown CD, Capri E, Gottesbüren BFF, Heimbach F, Lythgo CM, Schulz R and Streloke MSETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 304–320.

Boesten JJTI, Köpp H, Adriaanse PI, Brock TCM and Forbes VE 2007. Conceptual model for improving the link between exposure and effects in the aquatic risk assessment of pesticides. Ecotoxicology and Environmental Safety, 66, 291–308.

Brock TCM and Budde BJ, 1994. On the choice of structural parameters to indicate responses of freshwater ecosystems to pesticide stress. In: Freshwater Field Tests for Hazard Assessment of Chemicals. Eds Hill IA, Heimbach F, Leeuwangh P and Matthiesen P. Lewis Publishers, Michigan, USA, 19–56.

Brock TCM and Van Wijngaarden RPA, 2012. Acute toxicity tests with *Daphnia magna*, *Americamysis bahia*, *Chironomus riparius* and *Gammarus pulex* and implications of new requirements for the aquatic effect assessment of insecticides. Environmental Science and Pollution Research, 19, 3610–3618.

Brock TCM, Arts GHP, Maltby L and Van den Brink PJ, 2006. Aquatic Risks of Pesticides, Ecological Protection Goals and Common Aims in European Union Legislation. Integrated Environmental Assessment and Management, 2, E20–E46.

Brock TCM, Maltby L, Hickey CW, Chapman J and Solomon K, 2008. Spatial extrapolation in ecological effect assessment of chemicals. In: Extrapolation Practice for ecotoxicological effect characterization of chemicals. Eds Solomon KR, Brock TCM, De Zwart D, Dyer SD, Posthuma L, Richards SM, Sanderson H, Sibley PK and Van den Brink PJ. SETAC Press & CRC Press, Boca Raton, FL, USA, 223–256.

Brock TCM, Roessink I, Belgers JDM, Bransen F and Maund SJ, 2009. Impact of a benzoyl urea insecticide on aquatic macro-invertebrates in ditch mesocosms with and without non-sprayed sections. Environmental Toxicology and Chemistry, 28, 2191–2205.

Brock TCM, Alix A, Brown CD, Capri E, Gottesbüren BFF, Heimbach F, Lythgo CM, Schulz R and Streloke M (Eds), 2010a. Linking aquatic exposure and effects: risk assessment of pesticides. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 398 pp.

Brock TCM, Belgers JDM, Roessink I, Cuppen JGM and Maund SJ, 2010b. Macroinvertebrate responses to insecticide application between sprayed and adjacent non-sprayed ditch sections of different sizes. Environmental Toxicology and Chemistry, 29, 1994–2008.

Brock T, Arts G, Belgers D and Van Rhenen-Kersten C, 2010c. Ecological characterization of drainage ditches in the Netherlands to evaluate pesticide stress. In: Linking aquatic exposure and effects: risk assessment of pesticides. Eds Brock TCM, Alix A, Brown CD, Capri E, Gottesbüren BFF, Heimbach F, Lythgo CM, Schulz R and Streloke M. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 269–287.

Brock TCM, Arts GHP, Ten Hulscher TEM, De Jong FMW, Luttik R, Roex EWM, Smit CE and Van Vliet PJM, 2011. Aquatic effect assessment for plant protection products: a Dutch proposal that addresses the requirements of the Plant Protection Product Regulation and Water Framework Directive. Alterra Report No 2235.

Brown K, Tomlinson J, Duncan J, Hinchcliffe A and Palmquist K, 2009. Critical comparison of available and potential higher tier testing approaches for the risk assessment of plant protection products, considering at least field and semi-field experimental designs, extrapolation from dose–response relationships, and increased dosages (aquatic and terrestrial). Final report: CFT/EFSA/PPR/2008/1: Lot 4. http://www.efsa.europa.eu/en/supporting/doc/16e.pdf

Bundschuh M, Zubrod JP, Kosol S, Maltby L, Stang C, Duester L and Schulz R, 2011. Fungal composition on leaves explains pollutant-mediated indirect effects on amphipod feeding. Aquatic Toxicology, 104, 32–37.

Campbell PJ, Arnold DJS, Brock TCM, Grandy NJ, Heger W, Heimbach F, Maund SJ and Streloke M, 1999. Guidance document on higher-tier aquatic risk assessment for pesticides (HARAP). Brussels (BE): SETAC-Europe, 179 pp.

Caquet T, Lagadic L and Sheffield SR, 2000. Mesocosms in ecotoxicology. 1. Outdoor aquatic systems. Reviews of Environmental Contamination and Toxicology, 165, 1–38.

Caquet C, Hanson M, Roucaute M, Graham D and Lagadic L, 2007. Influence of isolation on the recovery of pond mesocosms from the application of an insecticide. II Benthic macroinvertebrate responses. Environmental Toxicology and Chemistry, 26, 1280–1290.

Carsel RF, Imhoff JC, Hummel PR, Cheplick JM and Donigian Jr AS, 1995. PRZM-3. A model for predicting pesticide and nitrogen fate in the crop root and unsaturated soil zones. Users Manual for Release 3.0. National Exposure Research Laboratory, US Environmental Protection Agency, Athens, GA, USA.

Chirico N and Gramatica P, 2011. Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the Concordance Correlation Coefficient. Journal of Chemical Information and Modeling, 51, 2320–2335.

Chirico N and Gramatica P, 2012. Real External Predictivity of QSAR Models. Part 2. New inter-comparable thresholds for different validation criteria and the need for scatter plot inspection. Journal of Chemical Information and Modeling, 52, 2044–2058.

Chow VT, 1959. Open-channel hydraulics. McGraw-Hill Book Company Inc., USA, 680 pp.

Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, pp. 1–84.

Commission Regulation (EU) No 284/2013 of 1 March 2013 setting out the data requirements for plant protection products, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, p. 85–152.

Consonni et aV, Ballabio D and Todeschini R, 2009. Comments on the definition of the Q2 parameter for QSAR validation. Journal of Chemical Information and Modeling, 49, 1668–1678.

Coors A and Frische T, 2011. Predicting the aquatic toxicity of commercial pesticide mixtures. Environmental Sciences Europe, 23.

Coors A, Kuckelkorn J, Mammers-Wirtz and Strauss T, 2006. Application of in-situ bioassays with macrophytes in aquatic mesocosm studies. Ecotoxicology, 15, 583–591.

Council Directive 91/414/EEC concerning the placing of plant protection products on the market of 15 July 1991, OJ L 230, 19.8.1991, p. 1.

Creton S, Clook M and Wheeler JR, in preparation. Application of the threshold approach for acute fish toxicity testing to plant protection products: a proposed framework.

Creton S, Weltje L, Hobson H and Wheeler JR, 2013. Reducing the number of fish in bioconcentration studies for plant protection products by reducing the number of test concentrations. Chemosphere, 90, 1300–1304.

Crossland NO, Heimbach F, Hill IR, Boudou A, Leeuwangh P, Matthiessen P and Persoone G, 1993. European Workshop on Freshwater Field tests (EWOFFT), Summary and recommendations. Workshop held in Potsdam, Germany, June 25 –26, 1992.

Dabrowski JM, Peall SKC, Reinecke AJ, Liess M and Schulz R, 2001. Runoff-related pesticide input into the Lourens river, South Africa: Basic data for exposure assessment and risk mitigation at the catchment scale. Water, Air and Soil Pollution, 135, 265–283.

De Jong FMW, Brock TCM, Foekema EM and Leeuwangh P, 2008. Guidance for summarizing and evaluating aquatic micro- and mesocosm studies. RIVM Report 601506009, Bilthoven, the Netherlands, 59 pp.

De Lange M, Sala S, Vighi M and Faber J, 2010. A framework for applying ecological vulnerability in risk assessment. 20th Annual Meeting SETAC Europe (Society of Environmental Toxicology and Chemistry). 23–27 May 2010, Seville, Spain.

De Laender F, Van Sprang P and Janssen C, 2013. A re-evaluation of fifteen years of European Risk Assessment using effect models. Environmental Toxicology and Chemistry, 32, 594–601.

Diepens NJ, Arts GHP, Brock TCM, Smidt H, Van den Brink PJ, Van den Heuvel-Greve MJ and Koelmans AA, 2013. Sediment toxicity testing of organic chemicals in the context of prospective RA: A review. Critical Reviews in Environmental Science and Technology (accepted).

Dijksterhuis J, Van Doorn T, Samson R and Postma J, 2011. Effects of seven fungicides on non-target aquatic fungi. Water, Air and Soil Pollution, 222, 421–425.

Dimitrov SD, Dimitrova GD, Pavlov TS, Dimitrova N, Patlewicz GY, Niemela J and Mekenyan OG, 2005. A stepwise approach for defining the applicability domain of SAR and QSAR models. Journal of Chemical Information and Modeling, 45, 839–849.

Duncan J, Hinchcliffe A and Palmquist K, 2009. Evidence of potential long term effects in (aquatic and terrestrial) invertebrates after short term pulsed exposure, Literature reviews on ecotoxicology of chemicals with special focus on plant protection products. Reference: FT/EFSA/PPR/2008/01. http://www.efsa.europa.eu/en/supporting/pub/17e.htm

EC (European Commission), 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. OJ L 327/1, 22.12.2000, pp. 1–72.

EC (European Commission), 2002a. Guidance Document on Aquatic Ecotoxicology in the context of the Directive 91/414/EEC (SANCO/3268/2001) rev.4 final, 17.11.2002, pp. 1–62.

EC (European Commission), 2002b. Guidance Document on Terrestrial Ecotoxicology in the context of the Directive 91/414/EEC (SANCO/10329/2002, rev.2 final, 17.10.2002, pp. 1–39.

EC (European Commission), 2002c. Guidance Document on Risk Assessment for Birds and Mammals under Council Directive 91/414/EEC. SANCO/4145/2000—final 25 September 2002, p. 74.[58]

EC (European Commission), 2003. Technical Guidance Document on risk assessment in support of Commission Directive 93/67/EEC on risk assessment for new notified substances and Commission Regulation (EC) No 1488/94 on risk assessment for existing substances and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market.

EC (European Commission), 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH),establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. OJ L 136/3, 29.5.2007, pp. 1–287.

EC (European Commission), 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. EN OJ L 353/1, 31.12.2008, pp. 1–1355.

EC (European Commission), 2009. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L 309/1, 24.11.2009, pp. 1–50.

EC (European Commission), 2011a. Regulation (EC) No 546/2011 of 10 June 2011 implementing Regulation (EC) No 1107/2009 of the European Parliament and of the Council as regards uniform principles for evaluation and authorisation of plant protection products.

EC (European Commission), 2011b. Technical Guidance for Deriving Environmental Quality Standards, Guidance Document No: 27 under the Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Technical Report 2011-055.

EC (European Commission), 2011c. Guidance Document, Guidelines on comparability, extrapolation, group tolerances and data requirements for setting MRLs. SANCO Document 7525/VI/95-rev.9, March 2011.

EC (European Commission), 2012. Guidance Document on the assessment of the equivalence of technical materials of substances regulated under Regulation (EC) No 1107/2009. SANCO/10597/2003 rev. 10.1, 13 July 2012.

ECB (European Chemicals Bureau), 2005. Scoping study on the development of a technical guidance document on information requirements on intrinsic properties of substances. Report prepared by CEFIC, DK-EPA, Environmental Agency of Wales and England, ECETOC, INERIS, KemI and TNO. European Chemicals Bureau, Joint Research Centre, European Commission, Ispra, Italy.

ECHA (European Chemicals Agency), 2008. Guidance on information requirements and chemical safety assessment. Chapter R.7b: Endpoint specific guidance. Version 1.1. Helsinki, Finland: European Chemicals Agency. 234 p.

EEA (European Environment Agency), 2009. Small water bodies—Assessment of the status and threats of standing small water bodies. EEA/ADS/06/001—Water, Version 1.1.

EFSA (European Food Safety Authority), 2005a. Opinion of the Scientific Panel on Plant health, Plant Protection Products and their Residues on a request from the EFSA related to the evaluation of dimoxystrobin. The EFSA Journal 2005, 178, 45 pp. doi:10.2903/j.efsa.2005.178.

EFSA (European Food Safety Authority), 2005b. Opinion of the Scientific Panel on Plant health, Plant protection products and their Residues on a request from EFSA on the appropriateness of using the current FOCUS surface water scenarios for estimating exposure for risk assessment in aquatic

[58] Available at: http://ec.europa.eu/food/plant/protection/evaluation/guidance/wrkdoc19_en.pdf

ecotoxicology in the context of Council Directive 91/414/EEC. The EFSA Journal 2005, 145, 31 pp. doi:10.2903/j.efsa.2005.145.

EFSA (European Food Safety Authority), 2006a. Opinion of the Scientific Panel on Plant health, Plant Protection Products and their Residues on a request from the EFSA related to the assessment of the acute and chronic risk to aquatic organisms with regard to the possibility of lowering the assessment factor if additional species were tested. The EFSA Journal 2005, 301, 45 pp. doi:10.2903/j.efsa.2006.301.

EFSA (European Food Safety Authority), 2006b. Opinion of the Scientific Panel on Plant Health, Plant Protection Products and their Residues (PPR) on a request from EFSA related to the evaluation of pirimicarb. The EFSA Journal 2005, 240, 21 pp. doi:10.2903/j.efsa.2005.240.

EFSA (European Food Safety Authority), 2006c. Opinion of the Scientific Panel on Plant health, Plant Protection Products and their Residues on a request from the EFSA related to the aquatic risk assessment for cyprodinil and the use of a mesocosm study in particular. The EFSA Journal 2006, 329, 77 pp. doi:10.2903/j.efsa.2006.329.

EFSA (European Food Safety Authority), 2007a, Opinion of the Scientific Panel on Plant protection products and their Residues on a request from the Commission related to the revision of Annexes II and III to Council Directive 91/414/EEC concerning the placing of plant protection products on the market—Ecotoxicological studies. The EFSA Journal 2007, 461, 44 pp. doi:10.2903/j.efsa.2007.461.

EFSA (European Food Safety Authority), 2007b. Opinion of the Scientific Panel on Plant protection products and their Residues on a request from the Commission related to the revision of Annexes II and III to Council Directive 91/414/EEC concerning the placing of plant protection products on the market—Fate and Behaviour in the Environment. The EFSA Journal 2007, 448, 17 pp. doi:10.2903/j.efsa.2007.448.

EFSA (European Food Safety Authority), 2007c. Opinion of the Scientific Panel on Plant protection products and their Residues on a request from the Commission on acute dietary intake assessment of pesticide residues in fruit and vegetables. The EFSA Journal 2007, 538, 88 pp. doi:10.2903/j.efsa.2007.538.

EFSA (European Food Safety Authority), 2007d. Opinion of the Scientific Panel on Plant protection products and their Residues on a request from the Commission on the risks associated with an increase of the MRL for dieldrin on courgettes. The EFSA Journal 2007, 554, 48 pp. doi:10.2903/j.efsa.2007.554.

EFSA (European Food Safety Authority), 2008. Scientific Opinion of the Panel on Plant protection products and their Residues (PPR) on the Science behind the Guidance Document on Risk Assessment for birds and mammals. The EFSA Journal 2008, 734, 181 pp. doi:10.2903/j.efsa.2008.734.

EFSA (European Food Safety Authority), 2009a. Scientific Opinion of the Panel on Plant Protection Products and their Residues on a request from EFSA updating the opinion related to Annex II and III: Ecotoxicological studies. The EFSA Journal 2009, 1165, 25 pp. doi:10.2903/j.efsa.2009.1165.

EFSA (European Food Safety Authority), 2009b. Outcome of the Public Consultation on the existing Guidance Documents on Aquatic and Terrestrial Ecotoxicology under Directive 91/414/EC. EFSA Journal 2009; 7(11):1375, 129 pp. doi:10.2903/j.efsa.2009.1375.

EFSA (European Food Safety Authority), 2009c. Guidance Document on Risk Assessment for Birds and Mammals on request of EFSA. EFSA Journal 2009; 7(12):1438, 358 pp. doi:10.2903/j.efsa.2009.1438.

EFSA (European Food Safety Authority), 2011. Submission of scientific peer-reviewed open literature for the approval of pesticide active substances under Regulation (EC) No 1107/2009. EFSA Journal 2011;9(2):2092, 49 pp. doi:10.2903/j.efsa.2011.2092.

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2010. Scientific Opinion on Statistical considerations for the safety evaluation of GMOs. EFSA Journal 2010;8(1):1250, 59 pp. doi:10.2903/j.efsa.2010.1250.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2010a. Scientific opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the Guidance Documents on Aquatic and Terrestrial Ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002). EFSA Journal 2010;8(10):1821, 55 pp. doi:10.2903/j.efsa.2010.1821.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2010b. Scientific Opinion on outline proposals for assessment of exposure of organisms to substances in soil. EFSA Journal 2010;8(1):1442, 38 pp. doi:10.2903/j.efsa.2010.1442.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2012a. Scientific Opinion on Evaluation of the Toxicological Relevance of Pesticide Metabolites for Dietary Risk Assessment. EFSA Journal 2012;10(07): 2799, 187 pp. doi:10.2903/j.efsa.2012.2799.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2012b. Scientific Opinion on the science behind the development of a Risk Assessment of Plant Protection Products on bees (Apis mellifera, Bombus spp. and solitary bees). EFSA Journal 2012;10(5):2668, 275 pp. doi:10.2903/j.efsa.2012.2668.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2012c. Scientific Opinion on the science behind the guidance for scenario selection and scenario parameterisation for predicting environmental concentrations in soil. EFSA Journal 2012;10(2):2562, 76 pp. doi:10.2903/j.efsa.2012.2562.

EFSA Scientific Committee, 2013. Scientific Opinion on the hazard assessment of endocrine disrupters: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. EFSA Journal 2013;11(3):3132, 84 pp. doi:10.2903/j.efsa.2013.3132.

Environment Canada, 2005. Guidance Document on Statistical Methods. EPS l/RM/46. Ottawa, ON, Canada; 2005.

EPPO, 2003. Environmental risk assessment scheme for plant protection products. Chapter 12: Non-target terrestrial higher plants. EPPO Bulletin, 33, 239–244.

Escher BI, Bramaz N, Richter M and Lienert J, 2006. Comparative ecotoxicological hazard assessment of beta-blockers and their human metabolites using a mode of action based test battery and a QSAR approach. Environmental Science and Technology, 40, 7402–7408.

FOCUS, 2000. FOCUS groundwater scenarios in the EU review of active substances. Report of the FOCUS Groundwater Scenarios Workgroup, EC Document Reference SANCO/321/2000 rev.2, 202 pp.

FOCUS, 2001. FOCUS Surface Water Scenarios in the EU Evaluation Process under 91/414/EEC. Report of the FOCUS Working Group on Surface Water Scenarios, EC Document Reference SANCO/4802/2001-rev.2, 245 pp.

FOCUS, 2007a. Landscape and mitigation factors in aquatic risk assessment. Volume 1. Extended Summary and Recommendations'. Report of the FOCUS Working Group on Landscape and Mitigation Factors in Ecological Risk Assessment, EC Document Reference SANCO/10422/2005 v2.0. 169 pp.

FOCUS, 2007b. Landscape And Mitigation Factors In Aquatic Risk Assessment. Volume 2. Detailed Technical Reviews. Report of the FOCUS Working Group on Landscape and Mitigation Factors in Ecological Risk Assessment, EC Document Reference, SANCO/10422/2005 v2.0. 436 pp.

FOCUS, 2008. Pesticides in Air: Considerations for Exposure Assessment. Report of the FOCUS Working Group on Pesticides in Air, EC Document Reference SANCO/10553/2006 Rev 2 June 2008. 327 pp.

FOCUS, 2009. Assessing Potential for Movement of Active Substances and their Metabolites to Ground Water in the EU. Report of the FOCUS Ground Water Work Group, EC Document Reference Sanco/13144/2010 version 1, 604 pp.

Foit K, Kaske O and Liess M, 2012. Competition increases toxicant sensitivity and delays the recovery of two interacting populations. Aquatic Toxicology, 106–107, 25–31.

Forbes VE and Calow P, 2002. Species Sensitivity Distributions revisited: a critical appraisal. Human and Ecological Risk Assessment, 8, 473–492.

Fryday S and Thompson H, 2012. Toxicity of pesticides to aquatic and terrestrial life stages of amphibians and occurrence, habitat use and exposure of amphibian species in agricultural environments. Supporting Publications 2012:EN-343. [348 pp.]. http://www.efsa.europa.eu/en/supporting/doc/343e.pdf

Gergs A, Classen S, Hommen U and Preuss TG, 2011. Identification of realistic worst case aquatic macroinvertebrate species for prospective risk assessment using the trait concept. Environmental Science and Pollution Research, 18, 1316–1323.

German Federal Environment Agency (UBA), 2013. Environmental Risk Assessment of Pesticide Mixtures—German Guidance for Implementation under Regulation 1107/2009/EC.

Giddings JM, Brock TCM, Heger W, Heimbach F, Maund SJ, Norman S, Ratte H-T, Schäfers C and Streloke M (eds). 2002. Community-level aquatic system studies-interpretation criteria. (CLASSIC) Pensacola (FL): SETAC 44 p.

Giddings J, Arts G and Hommen U, 2013. The Relative Sensitivity of Macrophyte and Algal Species to Herbicides and Fungicides: An Analysis Using Species Sensitivity Distributions. Integrated Environmental Assessment and Management, 9, 308–318.

Giesy JP and Graney RL, 1989. Recent developments in and intercomparisons of acute and chronic bioassays and bioindicators. Hydrobiologia 188/189, 21–60.

Heckman LH and Friberg N, 2005. Macroinvertebrate community response to pulse exposure with the insecticide lambda-cyhalothrin using in-stream mesocosms. Environmental Toxicology and Chemistry, 24, 582–590.

Heugens E, Hendriks A, Dekker T, Van Straalen NM and Admiraal W, 2001. A review of the effects of multiple stressors on aquatic organisms and analysis of uncertainty factors for use in risk assessment. Critical Reviews in Toxicology, 31, 247–284.

Holland PT, 1996. Glossary of terms relating to pesticides (IUPAC Recommendations 1996). Pure and Applied Chemistry, 68, 1167–1193.

Hommen U, Poethke HJ, Dulmer U, Ratte HT, 1993. Simulation-models to predict ecological risk of toxins in fresh-water systems. ICES Journal of Marine Science, 50, 337–347

Hull RN and Swanson S, 2006. Sequential analysis of lines of evidence—an advanced weight of evidence approach for ecological risk assessment. Integrated Environmental Assessment and Management, 2, 302–311.

Hutchinson TH, Shillabeer N, Winter MJ and Pickford DB, 2006. Acute and Chronic Effects of Carrier Solvents in Aquatic Organisms: a Critical Review. Aquatic Toxicology, 76, 69–92.

Jager T, Heugens EHW and Kooijman SALM, 2006. Making sense of ecotoxicological test results: towards application of process-based models. Ecotoxicology, 15, 305–314.

Jager T, Albert C, Preuss TG and Ashauer R, 2011. General unified threshold model of survival—a toxicokinetic-toxicodynamic framework for ecotoxicology. Environmental Science and Technology, 45, 2529–2540.

Jarvis NJ, 1994. The MACRO model (Version 3.1). Technical Description and sample simulations. Reports and Dissertations, 19, Department of Soil Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden, 51 pp.

Jarvis NJ, 2001. The MACRO model (version 4.3). Technical description.

Jaworska J, Nikolova-Jelizkova N, and Aldenberg T, 2005. Review of methods for QSAR applicability domain estimation by the training set. ATLA-Alternatives to laboratory animals, 33, 445–459.

Jongbloed RH, Traas, TP and Luttik R, 1996. A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators. II. Calculations for dichlorodiphenyltrichloroethane (DDT) and cadmium. Ecotoxicology and Environmental Safety, 34, 279–306.

Junghans M, Backhaus T, Faust M, Scholze M and Grimme LH, 2006. Application and validation of approaches for the predictive hazard assessment of realistic pesticide mixtures. Aquatic Toxicology, 76, 93–110.

Kattwinkel M, Römbke J and Liess M, 2012. Ecological recovery of populations of vulnerable species driving the risk assessment of pesticides. EFSA Supporting Publications 2012:EN-338. [98 pp.]. http://www.efsa.europa.eu/en/supporting/doc/338e.pdf

Klok C, de Vries P, Jongbloed R and Tamis J, 2012. Literature review on the sensitivity and exposure of marine and estuarine organisms to pesticides in comparison to corresponding fresh water species. Supporting Publications 2012:EN-357. [157 pp.].
http://www.efsa.europa.eu/en/supporting/doc/357e.pdf

Knillmann S, Stampfli NC, Beketov MA and Liess M, 2012a. Intraspecific competition increases toxicant effects in outdoor pond microcosms. Ecotoxicology, 21, 1857–1866.

Knillmann S, Stampfli NC, Noskov YA, Beketov MA and Liess M, 2012b. Interspecific competition delays recovery of daphnia spp. Populations from pesticide stress. Ecotoxicology, 21, 1039–1049.

Kosol S, 2011. The effects of pesticides on aquatic hyphomycet community structure and functioning—implications for ecological risk assessment. Thesis University of Sheffield.

Kortenkamp A, Backhaus T and Faust M, 2009. State of the art report on mixture toxicity. Final report, study Contract number 070307/2007/485103/ETU/D.1. European Commission, Brussels, Belgium, 391 pp.

Kühne R, Ebert R-U, von der Ohe PC, Ulrich N, Brack W and Schüürmann G, 2013. Read-across prediction of the acute toxicity of organic compounds toward the water flea *Daphnia magna*. Molecular Informatics, 32, 108–120.

Lahr J, 1997. Ecotoxicology of organisms adapted to life in temporary freshwater ponds in arid and semi-arid regions. Archives of Environmental Contamination and Toxicology, 32, 50–57.

Leeds-Harrison PB, Shipway CJP, Jarvis NJ and Youngs EG, 1986. The influence of soil macroporosity on water retention, transmission and drainage in a clay soil. Soil Use and Management, 2, 47–50

Liess M, 2002. Population response to toxicants is altered by intraspecific interaction. Environmental Toxicology and Chemistry, 21, 138–142.

Liess M and Beketov M, 2011. Traits and stress: keys to identify community effects of low levels of toxicants in test systems. Ecotoxicology, 20, 1328–1340.

Liess M and Beketov M, 2012. Rebuttal related to 'Traits and stress: keys to identify community effects of low levels of toxicants in test systems, Liess nad Beketov (2011). Ecotoxicology, 21, 300–303.

Liess M and Schulz R, 1999. Linking insecticide contamination and population response in an agricultural stream. Environmental Toxicology and Chemistry, 18, 1948–1955.

Liess M and von der Ohe PC, 2005. Analysing effects of pesticides on invertebrate communities in streams. Environmental Toxicology and Chemistry, 24, 954–965.

Liess M, Brown C, Dohmen P, Duquesne S, Hart A, Heimbach F, Kreuger J, Lagadic L, Maund S, Reinert W, Streloke M and Tarazona, J, 2005. Effects of pesticides in the field—EPIF, Brussels, Belgium, SETAC Press, 136.

Liess M, Pieters B and Duquesne S, 2006. Long-term signal of population disturbance after pulse exposure to an insecticide—rapid recovery of abundance, persistent alteration of structure. Environmental Toxicology and Chemistry, 25, 1326–1331.

Liess M, Schäfer R and Schriever C, 2008. The footprint of pesticide stress in communities—species traits reveal community effects of toxicants. Science of the Total Environment, 406, 484–490.

Lin LI, 1989. A concordance correlation coefficient to evaluation reproducibility. Biometrics, 45, 255–268.

Lin R, Buijse L, Dohmen P, Kosol S, Maltby L, Rocha Dimitrov M, Roessink I, Sinkeldam JA, Smidt H, Van Wijngaarden RPA and Brock TCM, 2012. Effects of the fungicide metiram in outdoor freshwater microcosms: Responses of invertebrates, primary producers and microbes. Ecotoxicology, 21, 1550–1569.

Luttik R, 2003. Risk Assessment Scheme for the Impact of Plant Protection Products on Birds and Mammals. Thesis, Leiden University, The Netherlands.

Luttik R, Hart A, Roelofs W, Craig P and Mineau P, 2011. Variation in the level of protection afforded to birds and crustaceans exposed to different pesticides under standard risk assessment procedures, Integrated Environmental Assessment and Management, 7, 459–465.

Mackay D, 1982. Correlation of bioconcentration factors. Environmental Science and Technology, 16, 274–278.

Madelin R, 2004. The importance of scientific advice in the Community decision making process. Opening address to the Inaugural Joint Meeting of the members of the Non-Food Scientific Committees. Directorate General for Health and Consumer Protection, European Commission, Brussels.

Maltby L, 1992. Detritus Processing. In: The Rivers Handbook Vol. 1. Eds Calow P and Petts GE. Blackwell Scientific Publications, Oxford, UK, 331–353.

Maltby L, Blake N, Brock TCM and Van den Brink PJ, 2005. Insecticide species sensitivity distributions: the importance of test species selection and relevance to aquatic ecosystems. Environmental Toxicology and Chemistry, 24, 379–388.

Maltby L, Brock TCM and Van den Brink PJ, 2009. Fungicide risk assessment for aquatic ecosystems: Importance of interspecific variation, toxic mode of action and exposure regime. Environmental Science and Technology, 43, 7556–7563.

Maltby L, Arnold D, Arts G, Davies J, Heimbach F, Pickl C and Poulsen V (Eds), 2010. Aquatic macrophyte risk assessment for pesticides. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 140 pp.

McGill BJ, Enquist BJ, Weiher E and Westoby M, 2006. Rebuilding community ecology from functional traits. Trends in Ecology and Evolution, 21, 178–185.

Ministère des Relations Extérieures, Coopération et Développement, 1984. Mémento de l'agronome. République Française. 1604 pp.

Mohr S, Berghahn R, Feibicke M, Meinecke S, Ottenströer, Schmiedling I, Schmiediche R and Schmidt R, 2007. Effects of the herbicide metazachlor on macrophytes and ecosystem function in freshwater pond and stream mesocosms. Aquatic Toxicology, 82, 73–84.

Mohr S, Feibicke M, Berghahn R, Schmiediche R and Schmidt R, 2008. Response of plankton communities in freshwater pond and stream mesocosms to the herbicide metazachlor. Environmental Pollution, 152, 530–542.

Morgan MG, Dowlatabadi H, Henrion M, Keith D, Lempert R, McBride S, Small M and Wilbanks T (Eds), 2009. Best Practice Approaches for Characterizing, Communicating, and Incorporating Scientific Uncertainty in Climate Decision Making. U.S. Climate Change Science Program, Synthesis and Assessment Product 5.2, Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research, January 2009.

Munn S and Goumeou M, 2013. Key scientific issues relevant to the identification and characterisation of endocrine disrupting substances. Report of the Endocrine Disrupters Expert Advisory Group (endocrine disrupter EAG). European Commission—Joint Research Centre—Institute for Health and Consumer Protection, EUR 25919. Available from: http://ihcp.jrc.ec.europa.eu/our_activities/food-cons-prod/endocrine_disrupters/jrc-report-scientific-issues-identification-endocrine-disrupting-substances/at_multi_download/file?name=LBNA25919ENN.pdf

Muschal M and Warne MS, 2003. Risk posed by pesticides to aquatic organisms in rivers of Northern inland New South Wales, Australia. Human and Ecological Risk Assessment, 9, 1765–1787.

NAFTA (North American Free Trade Agreement), 2011. (Quantitative) Structure Activity Relationship ((Q)SAR) Guidance Document. NAFTA Technical working Group on Pesticides (TWG). Available from: http://www.epa.gov/oppfead1/international/naftatwg/guidance/qsar-guidance.pdf

Netzeva TI, Worth A, Aldenberg T, Benigni R, Cronin MT, Gramatica P, Jaworska JS, Kahn S, Klopman G, Marchant CA, Myatt G, Nikolova-Jeliazkova N, Patlewicz GY, Perkins R, Roberts D, Schultz T, Stanton DW, van de Sandt JJ, W Tong, Veith G and Yang C, 2005. Current status of methods for defining the applicability domain of (quantitative) structure–activity relationships. The report and recommendations of ECVAM Workshop 52. ATLA—Alternatives to Laboratory Animals, 33, 155–173.

Niemi GJ, De Vore P, Detenbeck N, Taylor D, Lima A and Pastor J, 1990. Overview of case studies on recovery of aquatic systems from disturbance. Environmental Management, 14, 571–587.

Nikolova N and Jaworska J, 2003. Approaches to measure chemical similarity—a review. QSAR and Combinatorial Science, 22, 1006–1026.

OECD (Organisation for Economic Cooperation and Development), 2000. Guidance document on aquatic toxicity testing of difficult substances and mixtures, OECD series on testing and assessment Number 23, ENV/JM/MONO(2000)6.

OECD (Organisation for Economic Cooperation and Development), 2006. Guidance Document on Simulated Freshwater Lentic Field tests (outdoor microcosms and mesocosms). Series on Testing and Assessment, No 53, ENV/JM/MONO(2006)17, OECD Environment Directorate, Paris, 37 pp.

OECD (Organisation for Economic Cooperation and Development), 2007. Guidance document on the validation of (Quantitative) Structure-Activity Relationship ((Q)SAR) models. ENV/JM/MONO(2007)2.

OECD (Organisation for Economic Cooperation and Development), 2009. Guidance Document for using the OECD (Q)SAR application toolbox to develop chemical categories according to the OECD Guidance on grouping of chemicals. Series on testing and assessment Number 102.

OECD (Organisation for Economic Cooperation and Development), 2012a. Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption. Series on Testing and Assessment, No. 150 (ENV/JM/MONO(2012)22).

OECD (Organisation for Economic Cooperation and Development), 2012b. Fish Toxicity Testing Framework. Series on Testing and Assessment No. 171. ENV/JM/MONO(2012)16.

O'Neill RV, Gardner RH, Barnthouse LW, Suter GW, Hildebrand SG and Gehrs CW, 1982. Ecosystem risk analysis: A new methodology. Environmental Toxicology and Chemistry, 1, 167–177.

Park RA, Clough JS and Wellman MC, 2008. AQUATOX: Modeling environmental fate and ecological effects in aquatic ecosystems. Ecological Modelling, 213, 1–15.

Porcelli C, Boriani E, Roncaglioni A, Chana A and Benfenati E, 2008. Regulatory perspectives in the use and validation of QSAR. A case study: DEMETRA model for *Daphnia* toxicity. Environmental Science and Technology, 42, 491–496.

Posthuma L, Suter GWI and Traas TP (Eds), 2002. Species sensitivity distributions in ecotoxicology. CRC Press, Lewis, Boca Raton, FL, USA, 587 pp.

Posthuma L, De Zwart D, Solomon KR and Brock TCM, 2008. Guidance on the application of extrapolation methods in ecological exposure and effects characterization of chemicals. In: Extrapolation practice for ecotoxicological effect characterization of chemicals. Eds Solomon KR, Brock TCM, De Zwart D, Dyer SD, Posthuma L, Richards SM, Sanderson H, Sibley PK and Van den Brink PJ. SETAC Press & CRC Press, Boca Raton, Fl, USA, 281–322.

Preuss TG, Hammers-Wirtz M and Ratte HT, 2010. The potential of individual based population models to extrapolate effects measured at standardized test conditions to relevant environmental conditions—an example for 3,4-dichloroaniline on *Daphnia magna.* Journal of Environmental Monitoring, 12, 2070–2079.

Relyea R and Hoverman J, 2006. Assessing the ecology in ecotoxicology: a review and synthesis in freshwater systems. Ecology Letters, 9, 1157–1171.

Reynaldi S, Meiser M and Liess M, 2011. Effects of the pyrethroid fenvalerate on the alarm response and on the vulnerability of the mosquito larva *Culex pipiens molestus* to the predator *Notonecta glauca*. Aquatic Toxicology, 104, 56–60.

Roessink I, Merga LB, Zweers HJ, Van den Brink PJ, 2013. The neonicotinoid imidacloprid shows high chronic toxicity to mayfly nymphs. Environmental Toxicology and Chemistry, 32, 1096–1100.

Roex EWM, Van Gestel CAM, Van Wezel AP and Van Straalen NM, 2000. Ratios between acute aquatic toxicity and effects on population growth rates in relation to toxicant mode of action. Environmental Toxicology and Chemistry, 19, 685–693.

Romijn CAFM, Luttik R, v.d. Meent D, Slooff W and Canton JH, 1993. Presentation of a general algorithm to include effect assessment on secondary poisoning in the derivation of environmental quality criteria. Part 1. Aquatic food chains. Ecotoxicology and Environmental Safety, 26, 61–85.

Romijn CAFM, Luttik R, and Canton JH, 1994. Presentation of a general algorithm to include effect assessment on secondary poisoning in the derivation of environmental quality criteria. 2. Terrestrial food chains. Ecotoxicology and Environmental Safety, 27, 107–127.

Rozman KK and Doull J, 2000. Dose and time as variables of toxicity. Toxicology, 144, 169–178.

Rubach MN, Baird DJ, Boerwinkel MC, Maund SJ, Roessink I and Van den Brink PJ, 2012. Species traits as predictors for intrinsic sensitivity of aquatic invertebrates to the insecticide chlorpyrifos. Ecotoxicology, 21, 2088–2101.

Russom CL, Bradbury SP, Broderius SJ, Hammermeister DE and Drummond RA, 1997. Predicting modes of toxic action from chemical structure: acute toxicity in the fathead minnow (*Pimephales promelas*). Environmental Toxicology and Chemistry, 16, 948–967.

Sanderson H, Laird B, Brain R, Wilson CJ and Solomon KR, 2009. Detectability of fifteen aquatic micro/mesocosms. Ecotoxicology, 18, 838–845.

Schäfer RB, Caquet T, Siimes K, Mueller R, Lagadic L and Liess M, 2007. Effects of pesticides on community structure and ecosystem functions in agricultural streams of three biogeographical regions in Europe. Science of the Total Environment, 382, 272–285.

Schäfer RB, von der Ohe PC, Rasmussen J, Kefford BJ, Beketov MA, Schulz R and Liess M, 2012. Thresholds for the effects of pesticides on invertebrate communities and leaf breakdown in stream ecosystems. Environmental Science and Technology, 46, 5134–5142.

SCHER (Scientific Committee on Health and Environmental Risks), SCCS (Scientific Committee on Consumer Safety), SCENIHR (Scientific Committee on Emerging and Newly Identified Health Risks), 2012. Opinion on the Toxicity and Assessment of Chemical Mixtures, 55 pp.

Schulz R and Liess M, 2000. Toxicity of fenvalerate to caddisfly larvae: chronic effects of 1- vs 10-h pulse-exposure with constant doses. Chemosphere, 41, 1511–1517.

Schüürmann G, Kühne R, Kleint F, Ebert R-U, Rothenbacher C and Herth P, 1997. A software system for automatic chemical property estimation from molecular structure, QSAR in Environmental Sciences VII. SETAC Press, Pensacola, FL, USA.

Schüürmann G, Ebert, RU, Chen JW, Wang B and Kühne R, 2008. External validation and prediction employing the predictive squared correlation coefficients Test Set Activity means vs Training Set Activity Mean. Journal of Chemical Information and Modeling, 48, 2140–2145.

Schüürmann G, Ebert R-U and Kühne R, 2011. Quantitative Read-across for Predicting the Acute Fish Toxicity of Organic Compounds. Environmental Science and Technology, 45, 4616–4622.

Shi LM, Fang H, Tong W, Wu J, Perkins R, Blair RM, Branham WS, Dial SL, Moland CL and Sheehan DM, 2001. QSAR models using a large diverse set of estrogens. Journal of Chemical Information and Computer Science, 41, 186–195.

Sijm DTHM, Van Wezel AP, Crommentuijn T, 2002. Environmental risk limits in the Netherlands. In: Species sensitivity distributions in ecotoxicology. Eds Posthuma L, Suter GWI and Traas TP. CRC Press, Lewis, Boca Raton, FL, USA, 221–253.

Sinclair CJ, 2009. Predicting the environmental fate and ecotoxicological and toxicological effects of pesticide transformation products. PhD Thesis. University of York, Environment Department.

Sinclair CJ and Boxall ABA, 2003. Assessing the ecotoxicity of pesticide transformation products, Environmental Science and Technology, 37, 4617–4625.

Solomon KR, Brock TCM, De Zwart D, Dyer SD, Posthuma L, Richards SM, Sanderson H, Sibley PK and Van den Brink PJ, (Eds), 2008. Extrapolation practice for ecotoxicological effect characterisation of chemicals. SETAC Press & CRC Press, Boca Raton, FL, USA, 380 pp.

Sokal RR and Rohlf FJ, 1995. Biometry, 3rd edn. W.H. Freeman, New York, USA.

Stäb JA, Traas TP, Stroomberg G, Van Kesteren J, Leonards P, VanHattum B, Brinkman UAT and Cofino WP, 1996. Determination of organotin compounds in the foodweb of a shallow freshwaterlake in The Netherlands. Archives of Environmental Contamination and Toxicology, 31, 319–328.

Stampfli NC, Knillmann S, Liess M and Beketov MA, 2011. Environmental context determines community sensitivity of freshwater zooplankton to pesticides. Aquatic Toxicology, 104, 116–124.

Sterling A, 2010. Keep it complex. Nature, 468, 1029–1031.

Suter GW, Traas TP and Posthuma L, 2002. Issues and practises in the derivation and use of species sensitivity distributions. In: Species sensitivity distributions in risk assessment. Eds Posthuma L, Traas TP and Suter GW. CRC Press, Boca Raton, FL, 473–474.

Ter Horst MMS, Adriaanse PI and Boesten JJTI, 2009. Mitigation of runoff in the FOCUS Surface Water Scenarios. Note of the fate group of the Environmental Risk Assessment team of Alterra on the interpretation of the mitigation of runoff in the FOCUS Landscape and Mitigation report (2007). Alterra rapport 1794, Wageningen.

Thorbek P, Forbes VE, Heimbach F, Hommen U, Thulke H-H, Van den Brink PJ, Wogram J and Grimm V (Eds), 2010. Ecological models for regulatory risk assessment of pesticides. Developing a strategy for the future. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA,125 pp.

Traas TP, Luttik R, and Jongbloed RH, 1996. A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators. I. Model description and uncertainty analysis. Ecotoxicology and Environmental Safety, 34, 264–278.

Traas TP, Janse JH, Van den Brink PJ, Brock TCM and Aldenberg T, 2004. A freshwater food web model for the combined effects of nutrients and insecticide stress and subsequent recovery. Environmental Toxicology and Chemistry, 23, 521–529.

Turner II BL, Kasperson RE, Matson PA, McCarthy JJ, Corell RW, Christensen L, Eckley N, Kaperson JX, Luers A, Martello ML, Polsky C, Pilsipher A and Schiller A, 2003. A framework for vulnerability analysis in sustainability science. Proceedings of the National Academy of Sciences of the USA, 100, 8074–8079.

US EPA (United States Environmental Protection Agency), 2008. Ecological Structure Activity Relationships. http://www.epa.gov/oppt/newchems/tools/21ecosar.htm, Washington DC, USA.

Van Beelen P, 2003. A review on the application of microbial toxicity tests for deriving sediment quality guidelines. Chemosphere, 53, 795–808.

Van den Brink PJ and Ter Braak CJF, 1998. Multivariate analysis of stress in experimental ecosystems by Principal Response Curves and similarity analysis. Aquatic Ecology, 32, 161–178.

Van den Brink PJ and Ter Braak CJF, 1999. Principal response curves: Analysis of time-dependent multivariate responses of biological communities to stress. Environmental Toxicology and Chemistry, 18, 138–148.

Van den Brink PJ and Ter Braak CJF, 2012. Response to 'Traits and stress: keys to identify community effects of low levels of toxicants in test systems' by Liess and Beketov (2011). Ecotoxicology, 21, 297–299.

Van den Brink PJ, Van Donk E, Gylstra R, Crum SJH and Brock TCM, 1995. Effects of chronic low concentrations of the pesticides chlorpyrifos and atrazine in indoor freshwater microcosms. Chemosphere, 31, 3181–3200.

Van den Brink PJ, Van Wijngaarden RPA, Lucassen WGH, Brock TCM and Leeuwangh P, 1996. Effects of the insecticide Durban 4®E (a.i. chlorpyrifos) in outdoor experimental ditches: II. Community responses and recovery. Environmental Toxicology and Chemistry, 15, 1143–1153.

Van den Brink PJ, Blake N, Brock TCM and Maltby L, 2006. Predictive value of species sensitivity distributions for effects of herbicides in freshwater ecosystems. Human and Ecological Risk Assessment, 12, 645–674.

Van den Brink PJ, Baveco JM, Verboom J and Heimbach F, 2007. An individual-based approach to model spatial population dynamics of invertebrates in aquatic ecosystems after pesticide contamination. Environmental Toxicology and Chemistry, 26, 10, 2226–2236.

Van der Hoeven N, 2008. Calculation of the minimum significant difference at the NOEC using a non-parametric test. Ecotoxicology and Environmental Safety 70, 61–66

Van Vlaardingen PLA, Traas TP, Wintersen AM and T. Aldenberg, 2004. ETX 2.0—A Program to Calculate Hazardous Concentrations and Fraction Affected, Based on Normally Distributed Toxicity Data, RIVM report 601501028/2004.

Van Wijngaarden RPA, Van den Brink PJ, Oude Voshaar JH and Leeuwangh P, 1995. Ordination techniques for analysing the responses of biological communities to toxic stress in experimental ecosystems. Ecotoxicology, 4, 61–77.

Van Wijngaarden RPA, Cuppen JGM, Arts GHP, Crum SHJ, Van den Hoorn MW, Van den Brink PJ and Brock TCM, 2004. Aquatic risk assessment of a realistic exposure to pesticides used in bulb crops: A microcosm study. Environmental Toxicology and Chemistry, 23, 1479–1498.

Van Wijngaarden RPA, Brock TCM and Maltby L, 2013. Calibration of Tier-2 assessments (geomean approach and SSD approach) for insecticides with results from model ecosystem experiments. Poster at SETAC 23rd Annual Meeting 2013, Glasgow.

Verhaar HJM, van Leeuwen CJ, and Hermens JLM, 1992. Classifying environmental pollutants. 1: Structure–activity relationships for prediction of aquatic toxicity. Chemosphere, 25, 471–491.

Von der Ohe PC and Liess M, 2004. Relative sensitivity distribution of aquatic invertebrates to organic and metal compounds. Environmental Toxicology and Chemistry, 23, 150–156.

Warne MStJ, 2003. A Review of the ecotoxicity of mixtures, approaches to, and recommendations for, their management. In: Proceedings of the 5th national workshop on the assessment of site contamination. Eds Langley A, Gilbey M and Kennedy B. National Environment Protection Council (NEPC), Adelaide, South Australia, 253–276. Available on http://www.ephc.gov.au/sites/default/files/ASC_WkshopPaper__19_Mix_Warne_Ecotoxicity_200301.pdf

Weltje L, Simpson P, Gross M, Crane M and Wheeler J, 2013. Comparative acute and chronic sensitivity of fish and amphibians: a critical review of data. Environmental Toxicology and Chemistry, 32, 984–994.

Williams JR, 1975. Sediment Yield Prediction with Universe Equation Using Runoff Energy Factor. In: Present and Prospective Technology for Predicting Sediment Yields and Sources. US Department of Agriculture, Washington, DC. ARS-S-40.

Willis KJ, Van den Brink PJ and Green JG, 2004. Seasonal variation in plankton community responses of mesocosms dosed with pentachlorophenol. Ecotoxicology, 13, 707–720.

Wogram J, 2010. Ecological characterization of small streams in northern and central Germany. In: Linking aquatic exposure and effects: risk assessment of pesticides. Eds Brock TCM, Alix A, Brown CD, Capri E, Gottesbüren BFF, Heimbach F, Lythgo CM, Schulz R, Streloke M. SETAC Press & CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 250–268.

Wogram J and Liess M, 2001. Rank ordering of macroinvertebrate species sensitivity to toxic compounds, by comparison with that of *Daphnia magna*. Bulletin of Environmental Contamination and Toxicology, 67, 360–367.

GLOSSARY AND ABBREVIATIONS

a.s.	active substance
AF	Assessment Factor
AGD	Aquatic Guidance Document
AMRAP	Aquatic Macrophyte Risk Assessment for Pesticides, SETAC Europe, 2009 2nd SETAC Europe Special Science Symposium, Brussels, Belgium, 2009–09–17/2009–09–18
AMPERE	Aquatic Mesocosms in Pesticide Registration in Europe: Recent Experiences (AMPERE) SETAC Europe Workshop, Leipzig, Germany, 24 – 25 April 2007
AUC	Area Under the Curve
BCF	Bioconcentration Factor
BMF	Biomagnification Factor
CA	Concentration Addition
CCC	Concordance Correlation Coefficient
CLASSIC	Community Level Aquatic System Studies Interpretation Criteria (CLASSIC). SETAC Europe/OECD/EC/BBA/UBA Workshop, 1999
DG SANCO	Directorate General for Health and Consumer Affairs
ECHA	European Chemicals Agency
EC_x	Concentration where x % effect was observed/calculated
EFSA	European Food Safety Authority
ELink	Linking Aquatic Exposure and Effects in the Registration Procedure of Plant Protection Products (Brock et al., 2010a)
EPPO	European and Mediterranean Plant Protection Organization
ERA	Environmental Risk Assessment
ERC	Ecotoxicologically Relevant Concentration
EQS	Environmental Quality Standard
ERO	Ecological Recovery Option
ETO	Ecological Threshold Option
ETR	Exposure-Toxicity Ratio
EU	European Union
Exposure profile	The course of time of the concentration on a relative concentration scale (an

	effect study is usually carried out at different concentration levels but with the same exposure profile).
FFLC test	Fish Full Life Cycle test
FOCUS	FOrum for the Co-ordination of pesticide fate models and their USe
GD	Guidance Document
HARAP	Higher-tier Aquatic Risk Assessment for Pesticides (HARAP), SETAC Europe workshop
HC_x	Hazardous concentration for x % of the species of a SSD
$LLHC_5$	Lower limit of the confidence interval of the hazardous concentration for 5 % of the species of a SSD
LOEC	Lowest Observed Effect Concentration
Metabolite	Any metabolite or a degradation product of an active substance, safener or synergist, formed either in organisms or in the environment (thus including also oxidation products which may have a larger molecular mass than the parent substance) (EFSA, 2012c).
MDD	Minimal Detectable Difference
MRR	Maximum Recommended Application Rate
NOEAEC	No Observed Ecologically Adverse Effect Concentration
NOEC	No Observed Effect Concentration
OECD	Organization for Economic Cooperation and Development
(Q)SAR	(Quantitative) Structure-Activity Relationship
PEC	Predicted Environmental Concentration
POM	Particulate Organic Matter
PPP	Plant Protection Product
PPR Panel	EFSA's Panel on Plant Protection Products and their Residues
PRC	Principle Response Curve
RA	Risk Assessment
RAC	Regulatory Acceptable Concentration
REACH	Registration, Evaluation, Authorisation and Restriction of Chemical substances
Ri	Reliability Index
SCFCAH	Standing Committee on the Food Chain and Animal Health

SED	sediment
SETAC	Society for Environmental Toxicology and Chemistry
SPG	Specific Protection Goal
SSD	Species Sensitivity Distribution
SW	Surface water
TK/TD	Toxicodynamics/toxicokinetic
TU	Toxic Unit
TWA	Time weighted average
WFD	Water Framework Directive

APPENDICES TO SCIENTIFIC OPINION

Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters

EFSA Panel on Plant Protection Products and their Residues (PPR)

European Food Safety Authority (EFSA), Parma, Italy

TABLE OF CONTENTS

APPENDIX A. ELEMENTS OF THE EXPOSURE ASSESSMENT GOALS RELATED TO THE CHOICES MADE IN THE $FOCUS_{SW}$ SCENARIOS

A1. Introduction

The aquatic risk assessment is the combination of the exposure and the effect assessments and there is considerable interaction between these assessments. The EFSA Panel on Plant Protection Products and their Residues (PPR Panel) (2010) indicated that the specification of the spatio-temporal statistical population of exposure concentrations together with the percentile to be taken from this spatio-temporal population are essential parts of the protection-goal dimensions because the risk is only assessed for the spatio-temporal variability of the systems that are included (so for the remaining systems it cannot be ruled out that high exposure concentrations leading to unacceptable effects will occur).

For this guidance document the PPR Panel assumes that the current exposure assessment procedure ($FOCUS_{SW}$ scenarios and models) will continue to be used and does not include further guidance for the exposure assessment. The EFSA PPR Panel does also not expect to evaluate the current exposure assessment (FOCUS, 2001, 2007a, b) in the coming years due to limited resources. As described above, the overall level of protection of aquatic organisms is determined by the combination of the specific protection goals for the organisms and the exposure assessment goals. So without description of the exposure assessment goals, the overall level of protection for aquatic organisms is undefined (EFSA, 2010). Therefore, we describe below the elements of the exposure assessment goals and add as far as possible the choices made by FOCUS (2001, 2007a, b).

The exposure assessment by FOCUS is based on diffuse sources of pollution that will occur if the plant protection product is applied and used following the rules of good agricultural practice (FOCUS, 2001, 2007a, b), so not including point sources resulting from inappropriate agricultural practices such as cleaning of spraying equipment and discharging the contaminated water directly into surface water systems. This is in line with Regulation (EC) No 1107/2009 which requires that plant protection products are authorised based on application consistent with good plant protection practice and having regard to realistic conditions of use.

Since around 2000, it has become common practice in the Standing Committee on the Food Chain and Animal Health (SCFCAH) to accept exposure assessments based on 90th percentile concentrations, as this is considered to be 'realistic worst case'. The definition of the exposure assessment goal then has to focus on the types of concentrations to be considered, e.g. a spray drift event will cause a much higher concentration in a shallow stream that is 1 cm deep than in a stream that is 30 cm deep. This specification is in the next sections split into (i) the spatial unit, (ii) the spatial statistical population of spatial units, and (iii) the temporal statistical population of concentrations. At the end the value of the percentile and its determination is discussed.

A2. Definition of the spatial unit

The definition of the spatial unit splits into two aspects: the type of spatial unit (e.g. edge-of-field water bodies that temporarily fall dry or that are permanent; macrophyte-dominated water bodies or all water bodies) and the size or area of this unit over which exposure concentrations may be averaged.

FOCUS (2001) developed 12 ditch and stream scenarios which all were 100 × 1 m and had a minimum water depth of 30 cm. FOCUS (2001) developed also three pond scenarios that were 30 × 30 m and at least 1 m deep. So these are all permanent water bodies, which is consistent with the effect assessment which is also based on permanent water bodies (see section 1.3.6).

FOCUS (2001) considered the concentration at the end of a ditch or stream that received spray drift, run-off or drainage from an adjacent field over a length of 100 m. So high local concentrations caused by outlets of individual drainpipes or by high local spray drift depositions (resulting, for example,

from vertical spray boom movements) were ignored. This approach is likely to provide concentrations that are close to the concentration averaged over a ditch or stream length of 100 m. This averaging length is related to the mobility of the organism whose effect is being assessed: in principle immobile organisms such as macrophytes would require averaging lengths as short as 1 m whereas for larger fish an averaging length of 1 000 m could be defensible. However, the specific protection goal for the macrophytes indicates that these organisms are protected at the population level. So for macrophytes this averaging over 100 m seems to be consistent with the proposed effect assessment.

FOCUS (2001) considered only average concentrations over the full surface area of the ponds.

A3. The spatial statistical population of the spatial units

After the spatial units have been defined, the spatial statistical population of these units can be defined. The first step is to specify the total area to be considered: for example the whole EU, one of the regulatory zones, a zone based on climate properties, a Member State or a major agricultural area within the EU such as the Po valley. This total area is related to the purpose of the regulatory decision making (e.g. EU registration, zonal registration or national registration).

We consider here only registration at EU level. Both FOCUS groundwater and FOCUS surface water scenarios were developed for some 10 locations distributed over the EU (then the EU-15; FOCUS, 2000, 2001) representing a range of climatic conditions. This was considered sufficient to identify a safe use of significant size. So the FOCUS surface water scenarios could be considered to apply to some 10 EU zones based on climate (and possibly soil) properties. The total area to be considered for each scenario would then be one of these 10 zones.

In the second step of the definition of the spatial statistical population it has to be decided whether all water systems in this area should be considered (i.e. a landscape-level approach) or only those adjacent to fields grown with the crop or crop group considered (i.e. the edge-of-field approach). FOCUS (2001) developed edge-of-field scenarios whereas FOCUS (2007a, b) described also (in great detail) methodologies for landscape-level exposure assessment. This guidance document deals only with edge-of-field water systems so we advise not using the landscape-level exposure assessment of FOCUS (2007a, b) in combination with this guidance document.

For the exposure assessment of edge-of-field surface waters, FOCUS (2001) assumed that the crop is grown and the substance is applied as close to the water as is possible considering good agricultural practice; this was supplemented by guidance by FOCUS (2007a, b) on emission reduction measures such as run-off buffer strips or spray-free zones. So the convention is that the crop is grown as close to the water as possible considering good agricultural practice.

FOCUS (2001) limited the population to water systems adjacent to fields treated with this active substance, so excluding systems adjacent to fields treated with other active substances.

FOCUS (2001) developed drainage scenarios for six locations (D1 to D6) and run-off scenarios for four locations (R1 to R4). The drainage scenarios receive only input from spray drift and drainpipes and the run-off scenarios receive only input from spray drift and run-off. The spray drift deposition values used by FOCUS were based on drift measurements that are downwind. FOCUS (2001) did not define the spatial populations on which the scenarios are based but the approach followed may be consistent with a statistical population that is further reduced based on the occurrence of entry routes (e.g. not considering surface water systems that are upwind during application or surface water systems that get only drainage or run-off inputs).

A4. Multi-year temporal statistical population of concentrations

The concentrations vary not only in space but also from year to year (see Figure A.1). So it has also to be defined which years are included in the statistical population and which are not. If a use in a single crop has to be evaluated, this is straightforward: i.e. include only the application years because these will nearly always generate the highest concentrations and because including all the years without applications in a rotation sequence does not make sense. However, if the same active substance is used in different crops in a crop rotation sequence, the exposure assessment becomes more complicated. It may of course happen that the peak concentration for crop no 2 becomes higher if crop no 1 is included in the exposure assessment (e.g. because drainpipe leaching resulting from application in crop no 1 leads to a higher concentration peak in the surface water due to spray drift at the time of application to crop no 2). Such combinations of uses in different crops are usually not included in exposure assessments at EU or national level. However, the definition of the temporal statistical population also has to be clear in case risk managers for some reason would want to assess such combinations in the future. Therefore, it is advisable to include in the definition of the temporal statistical population of concentrations also the use of the same a.s. in a crop rotation sequence.

Consider the following complicated but realistic example of a 4-year application sequence for a certain active substance:

- year 1: 1 kg/ha in maize and 0.5 kg/ha in carrots
- year 2: 0.7 kg/ha in sugar beet
- year 3: no applications
- year 4: no applications
- year 5: 1 kg/ha in maize and 0.5 kg/ha in carrots
- etc.

Let us assume that the effect assessment is based on annual peak concentrations. So each year in such a sequence generates one concentration and it has to be defined which concentration–year combinations are part of the temporal statistical population. FOCUS (2001) developed scenarios considering only one evaluation year in which a spray drift event takes place and the water system is also loaded with run-off and drainage. So years with low concentrations are not considered. This is consistent with including only the year with the highest concentration in the sequence in the statistical population (so only the years with the arrows in Figure A.1).

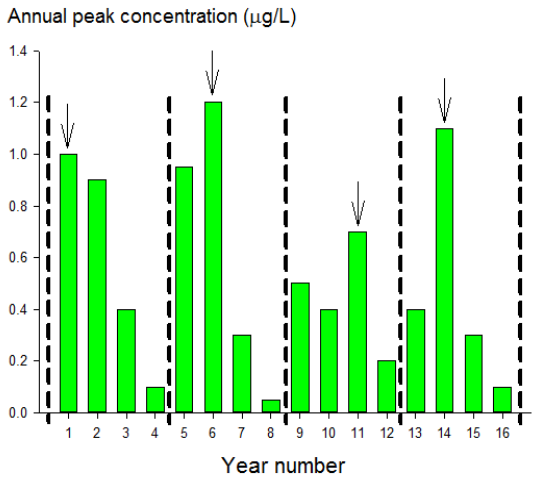


Figure A.1: Hypothetical sequence of annual peak concentrations for a period of 16 years with a four-year recurrence period of the application sequence (as in the example in the text). Dashed lines

indicate the 4-year periods and the arrows indicate the concentrations that are included in the temporal statistical population.

A5. Value of the percentile and its determination from the resulting combined spatio-temporal statistical population of concentrations

The concentration in the water systems is a function of both space and time. Let us assume for example that we have a population of 100 ditches and a 10-year time series of annual peak concentrations. Then we have a population of $100 \times 10 = 1000$ concentrations (henceforth called PECs). Let us assume that the 90th percentile is the target of the exposure assessment. The simplest procedure is just to derive the PEC considering all 1 000 values as one pool (i.e. one cumulative frequency distribution). Then the 90th percentile can be approximated by taking the 900th of the ranked PEC values. However, such an approach does not distinguish between space and time and has the consequence that the 100 PEC values above the 90th percentile PEC can be from all the years in 10 ditches thus accepting that the regulatory acceptable concentration (RAC) may be exceeded at 10 % of the locations in all years considered. But these 100 PEC values can also be from 1 year in 100 ditches, thus accepting that the RAC may be exceeded at all locations 10 % of the time (see Figure A.2 for an illustration of the possible combinations of spatial and temporal percentiles all giving an overall 90th percentile).

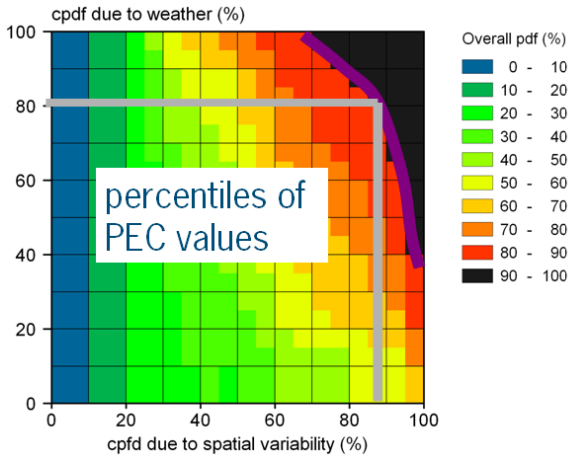


Figure A.2: Example of contour diagram of percentiles of exposure concentrations as a function of the spatial and temporal percentiles. The purple line shows all combinations that give an overall 90th percentile and the grey lines show a possible combination of a spatial and a temporal percentile giving an overall 90th percentile

An alternative approach would be to impose additional restrictions to selecting a percentile: for instance, do not assess the overall 90th percentile PEC but assess the 90th percentile PEC in space at the 50th percentile in time at all locations, thus accepting exceedance of the RAC for more than 50 % of the years only at 10 % of the locations. A priori it is unknown whether the overall 90th percentile PEC is lower than the 90th percentile PEC in space at the 50th percentile in time. FOCUS (2001) did not consider these aspects of the selection of a percentile.

Verdonck et al. (2003) showed that information on the species sensitivity variance may provide a rationale for such additional restrictions. However, these considerations are not easy to apply in a generic risk assessment. The PPR Panel considers that it is not feasible to include this aspect because there is in general no information available on the species composition of the surface waters at risk. The considerations of Verdonck et al. (2003) seem to be useful only for those compounds for which enough landscape-level and ecotoxicological information is available. For example, the consequences of such exceedances are considered more serious in a case in which all exceedances are concentrated

in a certain area and less serious in a case in which all exceedances are scattered over a whole regulatory zone. But this information is not available at a general level. Therefore the PPR Panel proposes to take simply the PEC corresponding to the required percentile of the cumulative probability density function without imposing any further restrictions with respect to the exceedance of the RAC in the space–time continuum.

The SCFCAH accepted FOCUS (2000) which based the assessment of leaching to groundwater on a 90th percentile. The SCFCAH also accepted FOCUS (2001, p. 7), which described the FOCUS surface water scenarios as 10 realistic worst-case scenarios which collectively represent agriculture in the EU. So FOCUS (2001) did not specify a value of the overall percentile, but used, for example for spray drift, 90th percentile values from the experimental drift dataset they selected. So in view of both FOCUS (2000) and FOCUS (2001) it seems that it was the intention of FOCUS (2001) to assess a 90th percentile. This is more or less supported by the statements on p. 109 of FOCUS (2001): (i) 'The various assumptions and 'worst-case' assessments summarised above show that, for many of the scenario factors that determine the magnitude and duration of pesticide residues in water bodies, a 90^{th+} percentile worst-case has been adopted.' and (ii) 'The highest PEC_{sw} estimates from the ten scenarios are likely to represent at least a 90th percentile worst-case for surface water exposures resulting from agricultural pesticide use within the European Union.'

As described before, EFSA PPR did not yet evaluate whether FOCUS (2001) achieved this 90th percentile protection level. Furthermore, risk managers have not yet taken a decision on the exposure assessment goals for surface water exposure assessment at the EU level.

Appendix B. Background to the procedure for partitioning of substances between water and sediment in the FOCUS$_{SW}$ step 2 exposure calculations

The effect of the two-thirds available/one-third non-available water compartments for sorption is shown for a strongly adsorbing, non-degrading compound (K_{oc} 10 000 L/kg) in Figure B.1. In the example simulation, a sequence of five applications, each of 1 kg/ha, is assumed (field crop, no run-off/drainage entry, no crop interception).

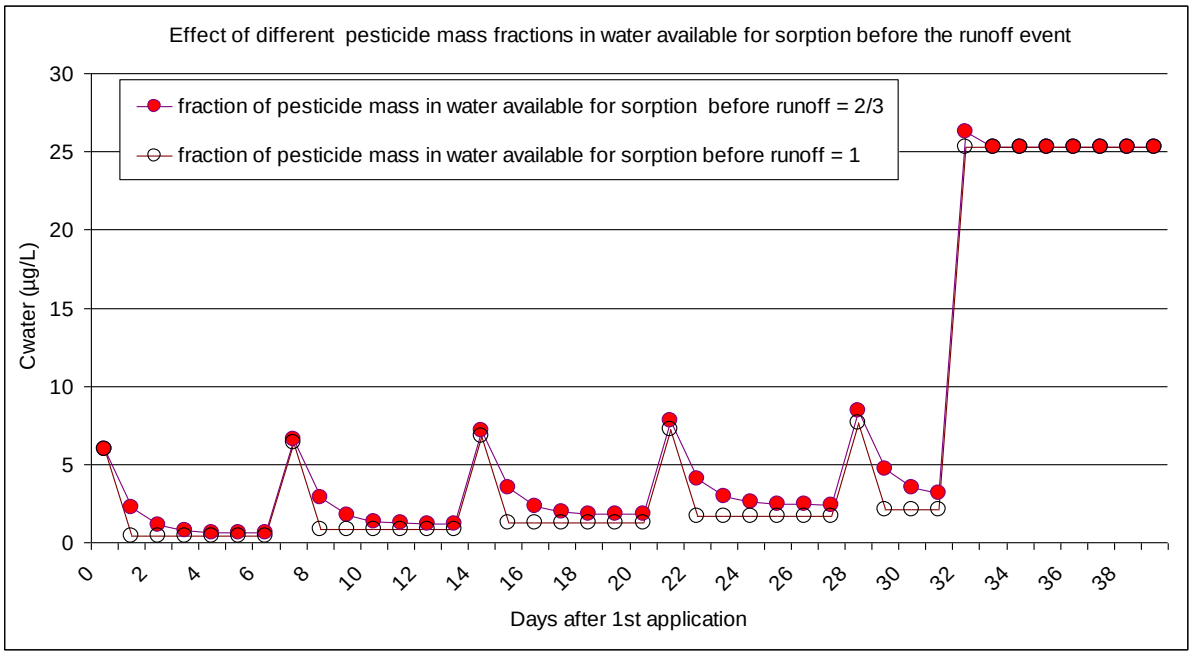


Figure B.1: Effect of the sediment/water distribution coefficient in the FOCUS step 2 model

As shown in Figure B.1 the concept of available/non-available water compartments delays the partitioning between water and sediment compared with the normal equilibrium approach.

After the run-off/drainage event (which occurred on day 32 in Figure B.1) the standard equilibrium equation between water and sediment is used by setting K, the fraction of PPP mass in water available for sorption, to 1. Consequently, the two lines in the diagram completely overlap after the run-off/drainage event.

The figure shows also that the steady-state value of the system with the special sediment/water distribution coefficient is higher than that of the normal equilibrium approach for the loads due to spray drift. This can be understood as follows. If F_{eq} is defined as the equilibrium fraction in the water phase as defined by the equations below and $F_{available}$ is defined as the fraction in the system in the available pool at steady state (so sum of available water and sorbed to sediment). In the steady-state situation, the fractions in the different compartments have to be as follows:

Available water: $F_{eq} \times F_{available}$ (consequence of definitions of F_{eq} and $F_{available}$)

Unavailable water: $0.5 \times F_{eq} \times F_{available}$ (because it is 50 % of the available water)

Sediment: $(1 - F_{eq}) \times F_{available}$ (consequence of definitions of F_{eq} and $F_{available}$)

The sum of the three compartments is $(1 + 0.5\ F_{eq})F_{available}$, which is, by definition, 1.0.

So the following is obtained:

$$F_{available} = 1/(1 + 0.5\ F_{eq})$$

Thus, the fraction in the water F_{water} in step 2 after spray applications can be calculated by:

$$F_{water} = 1.5\ F_{eq}/(1 + 0.5\ F_{eq}) = 3\ F_{eq}/(2 + F_{eq})$$

This can be illustrated with the first concentration peak in Figure B.1. F_{eq} for this system is 0.070, so the equilibrium concentration is $6 \times 0.070 = 0.42$ µg/L. The above equation then gives an F_{water} for this system of 0.101 so the equilibrium concentration in step 2 is $6 \times 0.101 = 0.61$ µg/L.

In FOCUS step 3 the partitioning between water and sediment is described with TOXSWA, which assumes a perfectly mixed water layer and diffusion into sediment. It is interesting to know to what extent the decline in water in step 2 is comparable to that in the equivalent stagnant TOXSWA system. The PPR Panel tested this by considering a system without degradation in water and sediment. For such a system, there is an analytical solution available (Equation 4.43 in Crank, 1967). Step 2 calculations were made for K_{oc} values of 10, 1 000 and 100 000 L/kg, a dosage of 1 kg/ha and an application that generated 12.1 % spray drift with no degradation in water and sediment. These were compared with the analytical solution assuming a dry bulk density of 0.8 kg/L, a porosity of 0.68, a tortuosity of 0.565, a diffusion coefficient in the liquid phase of 0.43 cm^2/d and the same dosage (and of course a water layer of 30 cm and an organic carbon content of 5 %). The porosity was calculated from the dry bulk density and the organic carbon content. The tortuosity was calculated using the same equation as in TOXSWA. The results in Figure B.2 indicate that for K_{oc} = 10 L/kg there is almost no substance going to the sediment in both systems. For K_{oc} = 1000 and 100 000 L/kg step 2 shows a much faster decline than the analytical solution.

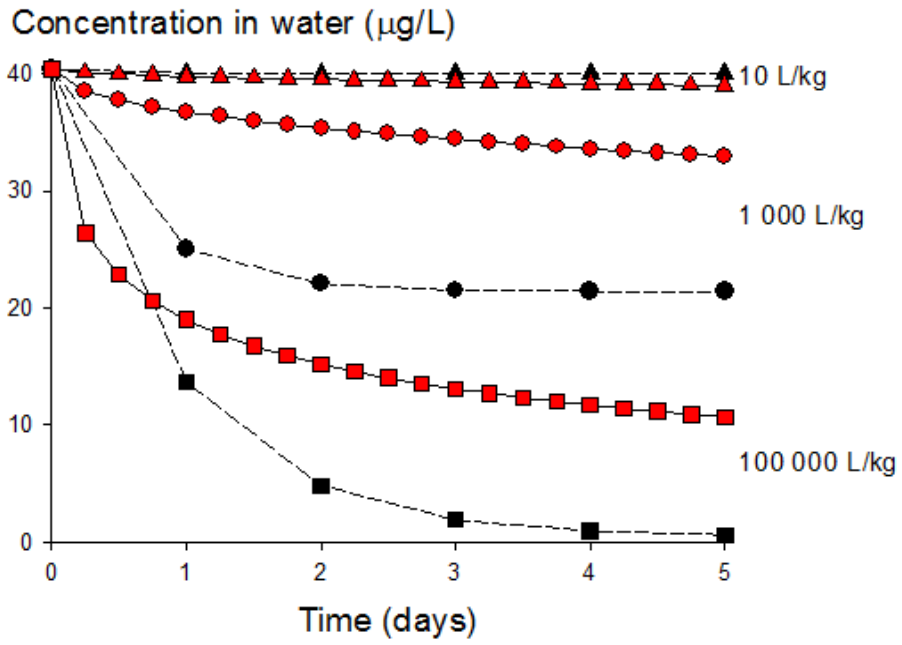


Figure B.2: Comparison between time course of the concentration in water calculated with (i) step 2 and (ii) the analytical solution for a perfectly mixed water layer with diffusion into the sediment (Equation 4.43 in Crank, 1967). Calculations are for a dosage of 1 kg/ha and an application that generated 12.1 % spray drift with no degradation in water and sediment. Dashed lines and black symbols are calculations with step 2 and solid lines and red symbols are calculations with the analytical solution for the three K_{oc} values indicated.

Appendix C. Comparison of acute rainbow trout toxicity with acute toxicity values for amphibian species

C.1 Introduction

Although the new data requirements (Commission Regulation (EU) No 283/2013[1]) do not request specific toxicity tests for amphibian species it is stated that available and relevant data, including data from the open literature for the active substance of concern, regarding the potential effects to amphibians shall be presented and taken into account in the risk assessment. It is not clear from this whether and when it would be necessary to test substances of concern on amphibians.

In the new data requirements (Commission Regulation (EU) No 283/2013) it is stated that a test on rainbow trout (*Oncorhynchus mykiss*) shall be carried out. In contrast with the old data requirements no species in addition to the rainbow trout is required. To determine whether standard tests with fish required for the dossier would be likely to cover the potential risk to amphibians present in the surface water, acute toxicity values for fish and amphibians have been compared.

This comparison used the data collected by Fryday and Thompson (2012) on amphibian species exposed in water. In particular, data used were generated in tests with an exposure duration of 96 hours and employing either a flow-through or a static-renewal exposure system.

Ideally the corresponding data for rainbow trout would have been taken from the dossiers, but the EFSA database on endpoints derived from the conclusions on pesticides included rainbow trout LC_{50} values for only five compounds that overlap with the Fryday and Thompson amphibian database. The second source for fish data was the Footprint IUPAC (International Union of Pure and Applied Chemistry) Pesticide Property Database (PPDB), and, where this database did not provide an LC_{50} for rainbow trout, safety data sheets from industry were used.

C.2 Data used for the comparison

In total, 253 data points for amphibian species with corresponding rainbow trout values were available, from tests on a plant protection product performed under either a flow-through or a static-renewal system. For 48 different species a toxicity test with a plant protection product was available (see Table C.3). Most of the tested species belong to the subclass of Anura (frogs and toads) and seven of the tested species to the subclass Caudata (salamanders or newts). 34 % of the tests were carried with *Xenopus laevis*, the African clawed frog. All individual values can be found in Fryday and Thompson (2012).

Tests are available for 60 different plant protection products: 7 fungicides, 19 herbicides, 32 insecticides, 1 plant growth regulator and 1 synergist (Table C.4). Only for two compounds was no LC_{50} value for rainbow trout available, and one study had only two days of exposure instead of the standard period of four days.

C.3 Results

Figure C.1 depicts the comparison of each amphibian toxicity value with the corresponding toxicity value for the rainbow trout (*O. mykiss*). The black line is the 1:1 line: the line indicating where toxicity to rainbow trout and amphibian would be exactly the same.

In 62 % of the cases the rainbow trout is more sensitive than the amphibian species (Table C.1 and C.2, points above the 1:1 line on Figure C1). The red line in Figure C1 represents an assessment factor of 100, i.e. where toxicity to an amphibian would be exactly 100 × higher than toxicity to the rainbow

[1] Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with the Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L 93, 3.4.2013, pp. 1–84.

trout. Only in 2 % of the cases is the amphibian test species more than a factor of 100 more sensitive than the rainbow trout (values below the red line in Figure C.1). Only in those cases would the LC_{50} for amphibians be lower than the RAC based on the rainbow trout.

The dataset of 253 tests with amphibians includes several life stages, e.g. tadpoles (including the Fryday and Thompson category 'larvae') and embryos. Repeating this analysis but splitting it by life stage (i.e. keeping embryos and larvae separate) gives a comparable result to the assessment on the whole dataset (see Table C.2). Therefore, the results are considered to be valid for both embryos and larvae. The amphibian toxicity values compared with rainbow trout values did not include adult life stages, and the level of protection of rainbow trout acute test data for adult amphibians therefore remains uncertain. However, it should be noted that for the purpose of the surface water risk assessment in field margins, fully aquatic life stages of amphibians (i.e. embryos and larvae) can be considered the most relevant.

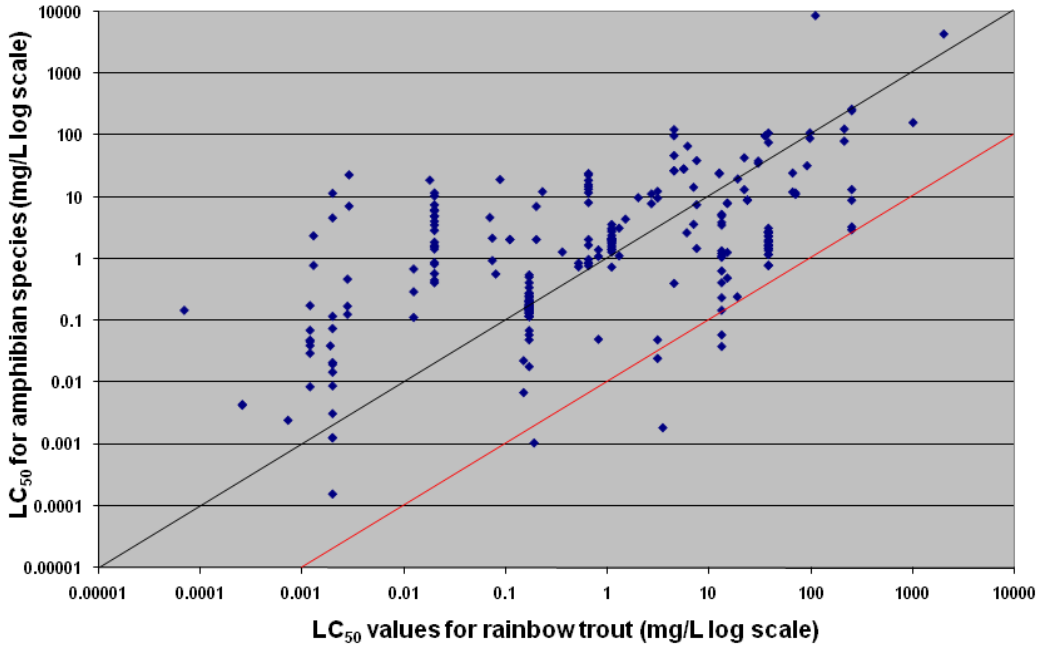


Figure C.1: Comparison of each amphibian toxicity value with the respective toxicity value for rainbow trout (*Oncorhynchus mykiss*). The black line is the 1:1 line, i.e. the line indicating that the outcome for rainbow trout and amphibians would be exactly the same. The red line considers the assessment factor of 100 applied in the acute risk assessment for fish.

Table C.1: Difference between amphibian species (embryos and tadpoles) and rainbow trout

Difference between amphibians and rainbow trout (*n* = 253)	**Amphibian species more sensitive than rainbow trout**	**Rainbow trout more sensitive than amphibian species**
More than factor of 1 000	0.4 %	2.8 %
Between 100 and 1 000 times	1.6 %	5.5 %
Between 10 and 100 times	15.4 %	18.2 %
Between 1 and 10 times	20.6 %	35.6 %
Less than a factor of 1	62.0 %	38.0 %

Table C.2: Differences between amphibian embryos or tadpoles and rainbow trout

Amphibians versus rainbow trout (embryos *n* = 54 and tadpoles *n* = 171)	Amphibian embryos more sensitive than rainbow trout	Amphibian tadpoles more sensitive than rainbow trout
> 1 000 times	0 %	1 %
100–1 000 times	4 %	4 %
10–100 times	13 %	11 %
1 and 10 times	11 %	37 %
< 1 times	72 %	47 %

Table C.3: Amphibian species included in the database/analysis and number of tests carried out with each species.

Amphibian species	Number of tests
Ambystoma gracile	3
Ambystoma laterale	1
Ambystoma maculatum	2
Ambystoma mexicanum	9
Ambystoma opacum	1
Bombina bombina	1
Bufo americanus	6
Bufo boreas	2
Bufo bufo gargarizans	9
Bufo japonicus formosus	6
Bufo melanostictus	1
Centrolene prosoblepon	1
Crinia insignifera	2
Cynops pyrrhogaster	7
Dendrosophus microcephalus	1
Engystomops pustulosus	1
Fejervarya limnocharis	1
Hyla japonica	6
Hyla versicolor	2
Hypsiboas crepitans	1
Hypsiboas pulchellus	2
Limnonectes limnocharis	1
Lithobates catesbeianus	1
Litoria moorei	2
Microhyla ornata	11
Notophthalmus viridescens	1
Physalaemus biligonigerus	1
Pseudacris crucifer	2
Pseudacris regilla	6
Rana cascadae	2
Rana catesbeiana	8
Rana clamitans	15
Rana cyanophlyctis	4
Rana hexadactyla	4
Rana limnocharis	1
Rana nigromaculata	7
Rana pipiens	9
Rana spinosa	1
Rana sylvatica	1

Amphibian species	Number of tests
Rhacophorus arboreus	4
Rhinella arenarum	12
Rhinella granulosa	1
Rhinella marina	1
Rhinella typhonius	1
Scinax nasicus	1
Scinax ruber	1
Silurana tropicalis	3
Xenopus laevis	87

Table C.4: Compounds tested for amphibians, fish toxicity values 96 hours LC_{50} in mg/L, and references (website where values were found including day of downloading).

Compound	LC_{50} (mg/L)	Remark	Day of downloading	References
2,4-D	250		11-11-2012	http://www.kellysolutions.com/erenewals/documentsubmit/KellyData%5COK%5Cpesticide%5CMSDS%5C42750 %5C42750-107 %5C42750-107_PD_2_5_8_2007_1_39_51_PM.pdf
Acephate	110		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Acetochlor	0.36		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Acrolein	0.15		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Acrylonitrile	70		11-11-2012	http://www.petrochemistry.net/ftp/pressroom/Microsoft%20Word%20-%20MSDS%20generic%203 %2008.pdf
Aldoxycarb	35.3		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Arsenous oxide	18.8		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Atrazine	4.5		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Azinphos-methyl	0.02		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Butachlor	0.52		11-11-2012	http://www.tlongagro.com/news_en/images/100728_1.pdf
Carbendazim	0.19		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Carbofuran	22		11-11-2012	http://www.mingdouchem.com/Enweb/UploadFiles/20100902102011404.pdf
Chlorimuron-ethyl	1000		11-11-2012	http://msds.dupont.com/msds/pdfs/EN/PEN_09004a3580169c5a.pdf
Chlorpyrifos	0.0013		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Copper sulphate	13.2		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Cypermethrin	0.0028		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
DDD	0.07		11-11-2012	http://webwiser.nlm.nih.gov/getSubstanceData.do;jsessionid=D71974FAF99B1421B5256A8607EFFA77?substanceID=30&displaySubstanceName=pp-TDE&UNNAID=&STCCID=&selectedDataMenuItemID=79
DDT	7		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Deltamethrin	0.00026		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Diazinon	3.1		10-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Dieldrin	0.0012		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Dimethoate	30.2		16-11-2012	EFSA database
Diuron	5.6		11-11-2012	http://nuturf.com.au/commerce/nuturf/msds133.pdf
Endosulfan	0.002		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/

Compound	LC_{50} (mg/L)	Remark	Day of downloading	References
Endrin	0.00073		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Esfenvalerate	0.00007		11-11-2012	http://msds.dupont.com/msds/pdfs/EN/PEN_09004a358060a40d.pdf
Fenitrothion	1.3		16-11-2012	EFSA database
Gamma-HCH, alpha-HCH, lindane	0.0029		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Glyphosate	38		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Glyphosate isopropyl-ammonium	7.5		11-11-2012	http://technical.nufarm.co.uk/documents/Herbicide/Safety/Clinic%20Ace.pdf
Imidacloprid	211		16-11-2012	EFSA database
Malathion	0.018		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Mancozeb	0.074		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Maneb	0.2		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Mefenacet	6	Carp[(a)]	9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Mercuric chloride	0.814		11-11-2012	http://scialert.net/fulltext/?doi=pjbs.2007.1098.1102
Myclobutanil	2		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Naphthalene	0.11		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Nicosulfuron	65.7		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Nonanoic acid	91		11-11-2012	http://apps.echa.europa.eu/registered/data/dossiers/DISS-9d8487fe-dd23-01ab-e044-00144f67d249/AGGR-3b0f8bc9-3bd9-4dac-b15e-7d14f920da40_DISS-9d8487fe-dd23-01ab-e044-00144f67d249.html
Paclobutrazol	23.6		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Paraquat	19		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Paraquat dichloride	15		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Parathion-methyl, methyl parathion	2.7		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Pentachloro-phenol	0.17		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Permethrin	0.0125		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Phosmet	0.23		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Piperonyl butoxide	6.12		11-11-2012	http://www.kellysolutions.com/erenewals/documentsubmit/KellyData%5COK%5Cpesticide%5CMSDS%5C655 %5C655-665 %5C655-665_Pyronyl_Ul_100_Concentrate_Prentox_9_22_2005_2_11_08_PM.pdf
Pirimicarb	79		16-11-2012	EFSA database
Prochloraz	1.5		16-11-2012	EFSA database
Profenofos	0.08		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Rotenone	0.0019		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Simetryn	7		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Sodium pentachloro-phenoxide	0.17	48 hours	11-11-2012	http://pmep.cce.cornell.edu/profiles/fung-nemat/febuconazole-sulfur/pentachlorophenol/prof-pentachlorophenol.html
Sulfometuron-methyl	12.5		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/
Temephos	3.49		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/

Compound	LC_{50} (mg/L)	Remark	Day of downloading	References
Thiobencarb	1.1		11-11-2012	http://www.bayercropscience.com.au/resources/uploads/msds/file7380.pdf?201211470908
Trichloro-acetic acid	2 000	*Lepomis macro-chirus*	11-11-2012	http://www.biovision.com/manuals/K860_MSDS.pdf?osCsid=sr2be5r0v8ocfvjbufqsj7bul2
Triclopyr-butotyl	0.65		11-11-2012	http://www.clarence.nsw.gov.au/content/uploads/grazon_dsmsd.pdf
Trifluralin	0.088		9-11-2012	http://sitem.herts.ac.uk/aeru/iupac/

(a): According to http://www.agropages.com/agrodata/Detail-891.htm the tested species is a carp.

Appendix D. Information on life cycle characteristics of aquatic organisms

This Appendix provides some indicative information in Table D.1 and Table D.2 on the life cycle characteristics of aquatic organisms available in the scientific literature and on the internet. This collection is not comprehensive and intended only to support, for example:.

- deciding if individuals may be exposed to repeated pulses of the active substance during their life span (section 9.1.3);
- deciding whether species are to be considered uni- or semivoltine (section 9.3.2.1).

This information should, however, be used with caution since the life cycle characteristics and overall life span may vary, e.g. depending on the climatic zone.

Table D.1: Generation time for various groups of aquatic organisms as derived from the PondFX Aquatic Life Database (www.ent.orst.edu/PondFX; http://new.freshwaterlife.org/freshwater-wiki/-/wiki/Main/Pond-FX) and Barnthouse (2004)

Taxon	Generation time in days [mean (range)]
Phytoplankton	1
Lemna	3
Rotifera	8 (6–35)
Cladocera	14
Copepoda	61 (14–73)
Oligochaeta	105 (51–730)
Amphipoda	73 (105–250)
Ostracoda	121 (51–362)
Gastropoda	513 (105–?)
Bivalvia	256 (105–?)
Coleoptera	(209–?)
Diptera	(81–503)
Ephemeroptera	(81–730)
Hemiptera	(81–503)
Trichoptera	(162–1 264)
Fish (short life cycle)	181
Fish (long life cycle)	1 673

Table D.2: Information on life cycle traits as extracted from the SPEAR documentation (http://www.systemecology.eu/spear/)

Taxon	Generation time (years)	Taxon	Generation time (years)
Acari	0.50	*Agabus paludosus*	1.00
Acentria	1.00	*Agabus* sp.	1.00
Acentria ephemerella	1.00	*Agabus sturmi*	1.00
Acilius	1.00	*Agabus subtilis*	1.00
Acilius canaliculatus	1.00	*Agabus uliginosus*	1.00
Acilius sp.	1.00	*Agabus undulatus*	1.00
Acilius sulcatus	1.00	*Agabus unguicularis*	1.00
Acroloxidae	1.00	*Agapetus*	0.63
Acroloxus	1.00	*Agapetus fuscipes*	0.50
Acroloxus lacustris	1.00	*Agapetus* sp.	0.75
Adicella	1.00	*Agraylea*	1.00
Adicella reducta	1.00	*Agraylea multipunctata*	1.00
Adicella sp.	1.00	*Agraylea sexmaculata*	1.00
Aedes	0.25	*Agraylea* sp.	1.00
Aedes sp.	0.25	*Agrypnia*	1.00
Aelosoma	0.50	*Agrypnia obsoleta*	1.00
Aelosomatidae	0.50	*Agrypnia pagetana*	1.00
Aeolosoma hemprichi	0.50	*Agrypnia* sp.	1.00
Aeshna	2.22	*Agrypnia varia*	1.00
Aeshna caerulea	3.00	*Alboglossiphonia*	1.00
Aeshna cyanea	2.00	*Alboglossiphonia heteroclita*	1.00
Aeshna grandis	3.00	*Allogamus*	1.00
Aeshna juncea	1.00	*Allogamus auricollis*	1.00
Aeshna mixta	2.00	*Allogamus* sp.	1.00
Aeshna sp.	2.30	*Allogamus uncatus*	1.00
Aeshnidae	2.00	Ameletidae	1.00
Aeshnidae sp.	2.00	*Ameletus*	1.00
Agabus	1.00	*Ameletus inopinatus*	1.00
Agabus affinis	1.00	*Ampharetidae*	0.50
Agabus arcticus	1.00	*Amphinemura*	1.00
Agabus biguttatus	1.00	*Amphinemura* sp.	1.00
Agabus bipustulatus	1.00	*Amphinemura sulcicollis*	1.00
Agabus chalconatus	1.00	*Amphipoda*	0.53
Agabus congener	1.00	*Anabolia*	1.00
Agabus conspersus	1.00	*Anabolia nervosa*	1.00
Agabus didymus	1.00	*Anacaena*	1.00
Agabus guttatus	1.00	*Anacaena bipustulata*	1.00
Agabus labiatus	1.00	*Anacaena globulus*	1.00
Agabus melanarius	1.00	*Anacaena limbata*	1.00
Agabus melanocornis	1.00	*Anacaena lutescens*	1.00
Agabus nebulosus	1.00	*Anacaena* sp.	1.00
Agabus neglectus	1.00	*Anax*	1.67
Agabus obscurus	1.00	*Anax imperator*	2.00

Anax parthenope	2.00	*Assiminea grayana*	1.70
Anax sp.	1.00	Assimineidae	1.70
Ancylidae	1.00	Astacidae	1.00
Ancylus	1.00	Athericidae	0.69
Ancylus fluviatilis	1.00	Athericidae sp.	0.75
Ancylus sp.	1.00	*Atherix*	0.67
Anisoptera	1.72	*Atherix ibis*	1.00
Anisus	1.00	*Atherix marginata*	0.50
Anisus leucostoma	1.00	*Atherix* sp.	0.50
Anisus sp.	1.00	*Athripsodes*	1.00
Anisus vortex	1.00	*Athripsodes albifrons*	1.00
Annitella	1.00	*Athripsodes aterrimus*	1.00
Annitella obscurata	1.00	*Athripsodes bilineatus*	1.00
Anodonta	1.00	*Athripsodes cinereus*	1.00
Anodonta anatina	1.00	*Athripsodes* sp.	1.00
Anodonta cygnea	1.00	*Atrichops*	0.50
Anodonta sp.	1.00	*Atrichops crassipes*	0.50
Anomalopterygella	1.00	*Atyaephyra*	1.00
Anomalopterygella chauvinia	1.00	*Atyaephyra desmaresti*	1.00
Aphanoneura	0.50	Atyidae	1.00
Aphelocheiridae	1.00	*Aulodrilus*	0.50
Aphelocheirus	1.00	*Aulodrilus pluriseta*	0.50
Aphelocheirus aestivalis	1.00	*Austropotamobius*	1.00
Aplexa	1.00	*Austropotamobius pallipes*	1.00
Aplexa hypnorum	1.00	Baetidae	0.60
Aquarius	0.50	*Baetis*	0.65
Aquarius paludum	0.50	*Baetis alpinus*	0.50
Arachnida	1.00	*Baetis buceratus*	0.50
Araneae	1.00	*Baetis digitatus*	0.50
Arctocorixa	0.67	*Baetis fuscatus*	0.50
Arctocorixa carinata	0.50	*Baetis muticus*	0.50
Arctocorixa germari	0.50	*Baetis niger*	0.50
Arctocorixa sp.	1.00	*Baetis rhodani*	0.50
Argulidae	0.40	*Baetis scambus*	2.00
Arguloidea	0.40	*Baetis* sp.	0.50
Argulus	0.40	*Baetis tracheatus*	0.60
Argulus foliaceus	0.40	*Baetis vernus*	0.50
Argyroneta	1.00	*Balanidae*	1.00
Argyroneta aquatica	1.00	*Balanus*	1.00
Arhynchobdellida	1.00	*Balanus balanoides*	1.00
Armiger	0.33	*Balanus improvisus*	1.00
Armiger crista	0.33	*Balanus* sp.	1.00
Asellidae	0.46	*Basommatophora*	0.94
Asellus	0.33	*Bathyomphalus*	1.00
Asellus aquaticus	0.33	*Bathyomphalus contortus*	1.00
Asellus meridianus	0.33	*Batracobdella*	1.00
Asellus sp.	0.33	*Batracobdella paludosa*	1.00
Assiminea	1.70	*Bdellocephala*	1.00

Bdellocephala punctata	1.00	*Bythinella* sp.	1.00
Beraea	1.00	Caenidae	0.71
Beraea pullata	1.00	*Caenis*	0.66
Beraeidae	1.00	*Caenis beskidensis*	0.50
Beraeodes	1.00	*Caenis horaria*	0.50
Beraeodes minutus	1.00	*Caenis luctuosa*	1.00
Berosus	1.00	*Caenis macrura*	0.50
Berosus affinis	1.00	*Caenis pseudorivulorum*	0.50
Berosus bicolor	1.00	*Caenis rivulorum*	1.00
Berosus luridus	1.00	*Caenis robusta*	0.50
Berosus signaticollis	1.00	*Caenis* sp.	0.80
Bezzia	0.33	*Callicorixa*	0.50
Bezzia sp.	0.33	*Callicorixa praeusta*	0.50
Bithynia	1.00	*Callicorixa wollastoni*	0.50
Bithynia leachi	1.00	Calopterygidae	1.78
Bithynia sp.	1.00	*Calopteryx*	1.78
Bithynia tentaculata	1.00	*Calopteryx* sp.	1.33
Bithyniidae	1.00	*Calopteryx splendens*	2.00
Bivalvia	0.90	*Calopteryx virgo*	2.00
Blephariceridae	1.00	Cambaridae	2.00
Blephariceridae sp.	1.00	*Canalipalpata*	0.50
Boreobdella	1.00	*Capnia*	1.00
Boreobdella verrucata	1.00	*Capnia bifrons*	1.00
Bosmina	0.33	*Capnia* sp.	1.00
Bosmina sp.	0.33	Capniidae	1.00
Bosminidae	0.33	*Carcinus*	1.00
Brachycentridae	1.00	*Carcinus maenas*	1.00
Brachycentrus	1.00	*Centroptilum*	0.50
Brachycentrus maculatum	1.00	*Centroptilum luteolum*	0.50
Brachycentrus maculatus	1.00	*Ceraclea*	1.00
Brachycentrus subnubilus	1.00	*Ceraclea albimacula*	1.00
Brachycercus	0.83	*Ceraclea dissimilis*	1.00
Brachycercus harrisella	1.00	*Ceraclea nigronervosa*	1.00
Brachycercus harrisellus	1.00	*Ceraclea senilis*	1.00
Brachycercus sp.	0.50	*Ceraclea* sp.	1.00
Brachyptera	1.00	Ceratopogonidae	0.33
Brachyptera risi	1.00	Ceratopogonidae sp.	0.33
Brachytron	1.00	*Cercyon*	1.00
Brachytron pratense	1.00	*Cercyon bifenestratus*	1.00
Branchiopoda	0.33	*Cercyon convexiusculus*	1.00
Branchiura	0.33	*Cercyon granarius*	1.00
Branchiura sowerbyi	0.33	*Cercyon marinus*	1.00
Brillia	0.40	*Cercyon* sp.	1.00
Brillia longifurca	0.40	*Cercyon sternalis*	1.00
Brychius	1.00	*Cercyon tristis*	1.00
Brychius elevatus	1.00	*Cercyon ustulatus*	1.00
Byrrhoidea	2.00	*Ceriagrion*	1.00
Bythinella	1.00	*Ceriagrion tenellum*	1.00

Chaetarthria	1.00	*Coelambus parallelogrammus*	1.00
Chaetarthria seminulum	1.00	*Coelostoma*	1.00
Chaetogaster	0.50	*Coelostoma orbiculare*	1.00
Chaetogaster diaphanus	0.50	*Coenagrion*	1.00
Chaetogaster diastrophus	0.50	*Coenagrion hastulatum*	1.00
Chaetogaster langi	0.50	*Coenagrion puella*	1.00
Chaetopterygini	1.00	*Coenagrion puella/pulchellum*	1.00
Chaetopterygini sp.	1.00	*Coenagrion pulchellum*	1.00
Chaetopteryx	1.00	*Coenagrion* sp.	1.00
Chaetopteryx major	1.00	Coenagrionidae	0.93
Chaetopteryx sp.	1.00	Coenagrionidae sp.	1.00
Chaetopteryx villosa	1.00	*Coleoptera*	1.07
Chalcolestes	1.00	*Colymbetes*	1.00
Chalcolestes viridis	1.00	*Colymbetes fuscus*	1.00
Chaoboridae	0.50	*Colymbetes kotulae*	1.00
Chaoboridae sp.	0.50	*Colymbetes* sp.	1.00
Chaoborus	0.50	Colymbetinae	1.00
Chaoborus sp.	0.50	Colymbetinae sp.	1.00
Cheumatopsyche	1.00	*Conchapelopia*	0.40
Cheumatopsyche lepida	1.00	*Conchapelopia melanops*	0.40
Chironomidae	0.33	*Conchapelopia* sp.	0.40
Chironomidae sp.	0.33	*Congeria*	1.00
Chironomidae sp. *rot*	0.33	*Congeria leucophaeata*	1.00
Chironomidae sp. weiß	0.33	*Copelatus*	1.00
Chironominae	0.40	*Copelatus haemorrhoidalis*	1.00
Chironomini	0.33	*Corbicula*	1.00
Chironomini sp.	0.33	*Corbicula fluminalis*	1.00
Chironomus	0.39	*Corbicula fluminea*	1.00
Chironomus plumosus Group	0.50	Corbiculidae	1.00
Chironomus sp.	0.33	*Cordulegaster*	3.50
Chironomus thummi Group	0.33	*Cordulegaster boltonii*	3.00
Chloroperla	1.00	*Cordulegaster* sp.	4.00
Chloroperla sp.	1.00	Cordulegastridae	3.50
Chloroperla torrentium	1.00	*Cordulia*	2.00
Chloroperlidae	1.00	*Cordulia aenea*	2.00
Chrysomelidae	1.00	Corduliidae	1.25
Chrysomelidae sp.	1.00	Corduliidae sp.	1.00
Chydoridae	0.33	*Corixa*	0.55
Cladocera	0.33	*Corixa affinis*	0.50
Cloeon	0.40	*Corixa dentipes*	0.50
Cloeon dipterum	0.30	*Corixa panzeri*	0.50
Cloeon simile	0.50	*Corixa punctata*	0.75
Coelambus	1.00	*Corixa* sp.	0.50
Coelambus confluens	1.00	Corixidae	0.90
Coelambus impressopunctatus	1.00	Corixidae sp.	0.90
Coelambus lautus	1.00	Corophiidae	0.33
Coelambus nigrolineatus	1.00	*Corophium*	0.33
Coelambus paralellogrammus	1.00	*Corophium curvispinum*	0.33

Corophium lacustre	0.33	*Diamesa*	0.33
Corophium multisetosum	0.33	*Diamesa* sp.	0.33
Corophium sp.	0.33	*Dicranota*	1.00
Corophium volutator	0.33	*Dicranota* sp.	1.00
Crambidae	1.00	*Dicrotendipes*	0.40
Crangon	1.00	*Dicrotendipes nervosus*	0.40
Crangon crangon	1.00	*Dicrotendipes* sp.	0.40
Crangonidae	1.00	*Dikerogammarus*	0.33
Crangonyctidae	0.33	*Dikerogammarus* sp.	0.33
Crangonyx	0.33	*Dikerogammarus villosus*	0.33
Crangonyx pseudogracilis	0.33	Diptera	0.43
Crenobia	1.00	*Diura*	1.00
Crenobia alpina	1.00	*Diura bicaudata*	1.00
Cricotopus	0.40	*Dixa*	0.50
Cricotopus sp.	0.40	*Dixa* sp.	0.50
Crunoecia	1.00	Dixidae	0.50
Crunoecia irrorata	1.00	Dixidae sp.	0.50
Cryptochironomus	0.40	Dolichopodidae	1.00
Cryptochironomus sp.	0.40	Dolichopodidae sp.	1.00
Culex	0.25	*Donacia*	1.00
Culex sp.	0.25	*Donacia dentata*	1.00
Culicidae	0.25	*Donacia semicuprea*	1.00
Culicidae sp.	0.25	*Donacia sparganii*	1.00
Culiseta	0.25	*Donacia versicolorea*	1.00
Culiseta sp.	0.25	*Donacia vulgaris*	1.00
Curculionidae	1.00	Donaciinae	1.00
Curculionidae sp.	1.00	Donaciinae sp.	1.00
Curculionoidea	1.00	*Dreissena*	1.00
Cybaeidae	1.00	*Dreissena polymorpha*	1.00
Cymatia	0.50	Dreissenidae	1.00
Cymatia bonsdorffi	0.50	*Drusus*	1.00
Cymatia coleoptrata	0.50	*Drusus annulatus*	1.00
Cymbiodyta	1.00	*Drusus biguttatus*	1.00
Cymbiodyta marginella	1.00	*Drusus* sp.	1.00
Cyrnus	1.00	Dryopidae	2.00
Cyrnus flavidus	1.00	Dryopidae sp.	2.00
Cyrnus trimaculatus	1.00	*Dryops*	2.00
Daphnia	0.33	*Dryops ernesti*	2.00
Daphnia sp.	0.33	*Dryops luridus*	2.00
Daphniidae	0.33	*Dryops similaris*	2.00
Decapoda	1.11	*Dryops* sp.	2.00
Demicryptochironomus	0.40	*Dryops striatellus*	2.00
Demicryptochironomus sp.	0.40	*Dugesia*	1.00
Dendrocoelidae	1.00	*Dugesia gonocephala*	1.00
Dendrocoelum	1.00	*Dugesia lugubris*	1.00
Dendrocoelum lacteum	1.00	*Dugesia lugubris/polychroa*	1.00
Deronectes	1.00	*Dugesia polychroa*	1.00
Deronectes latus	1.00	*Dugesia* sp.	1.00

Dugesia tigrina	1.00	*Enallagma*	1.00
Dugesiidae	1.00	*Enallagma cyathigerum*	1.00
Dytiscidae	1.00	*Enochrus*	1.00
Dytiscidae sp.	1.00	*Enochrus affinis*	1.00
Dytiscus	1.00	*Enochrus bicolor*	1.00
Dytiscus circumcinctus	1.00	*Enochrus coarctatus*	1.00
Dytiscus circumflexus	1.00	*Enochrus fuscipennis*	1.00
Dytiscus lapponicus	1.00	*Enochrus halophilus*	1.00
Dytiscus marginalis	1.00	*Enochrus isotae*	1.00
Dytiscus semisulcatus	1.00	*Enochrus melanocephalus*	1.00
Dytiscus sp.	1.00	*Enochrus ochropterus*	1.00
Ecdyonurus	0.50	*Enochrus quadripunctatus*	1.00
Ecdyonurus dispar	0.50	*Enochrus* sp.	1.00
Ecdyonurus sp.	0.50	*Enochrus testaceus*	1.00
Ecdyonurus torrentis	0.50	*Enoicyla*	1.00
Ecdyonurus venosus	0.50	*Enoicyla pusilla*	1.00
Echinogammarus	0.75	*Enoicyla* sp.	1.00
Echinogammarus berilloni	0.75	*Epeorus*	1.00
Echinogammarus sp.	0.75	*Epeorus sylvicola*	1.00
Ecnomidae	1.00	*Ephemera*	1.75
Ecnomus	1.00	*Ephemera danica*	2.00
Ecnomus tenellus	1.00	*Ephemera lineata*	2.00
Einfeldia	0.40	*Ephemera* sp.	1.00
Einfeldia sp.	0.40	*Ephemera vulgata*	2.00
Eiseniella	0.33	*Ephemerella*	1.00
Eiseniella tetraedra	0.33	*Ephemerella ignita*	1.00
Electrogena	0.50	*Ephemerella major*	1.00
Electrogena lateralis	0.50	*Ephemerella mucronota*	1.00
Electrogena quadrilineata	0.50	*Ephemerella notata*	1.00
Electrogena sp.	0.50	Ephemerellidae	0.92
Electrogena ujhelyii	0.50	Ephemeridae	1.75
Elmidae	2.00	Ephemeroptera	0.82
Elmidae sp.	2.00	*Ephoron*	1.00
Elmis	2.00	*Ephoron virgo*	1.00
Elmis aenea	2.00	*Epitheca*	1.00
Elmis latreillei	2.00	*Epitheca bimaculata*	1.00
Elmis maugetii	2.00	*Eriocheir*	1.00
Elmis sp.	2.00	*Eriocheir sinensis*	1.00
Elodes	1.00	*Eristalis*	0.25
Elodes elongata	1.00	*Eristalis* sp.	0.25
Elodes marginata	1.00	*Eristalomyia*	0.25
Elodes minuta	1.00	*Eristalomyia tenax*	0.25
Eloeophila	1.00	*Erpobdella*	1.00
Eloeophila sp.	1.00	*Erpobdella nigricollis*	1.00
Elophila	1.00	*Erpobdella octoculata*	1.00
Elophila nymphaeata	1.00	*Erpobdella* sp.	1.00
Empididae	0.50	*Erpobdella testacea*	1.00
Empididae sp.	0.50	*Erpobdella vilnensis*	1.00

Erpobdellidae	1.00	*Glyptotendipes*	0.40
Erythromma	1.00	*Glyptotendipes* Group A	0.50
Erythromma najas	1.00	*Glyptotendipes* sp.	0.40
Erythromma viridulum	1.00	*Goera*	1.00
Eukiefferiella	0.40	*Goera pilosa*	1.00
Eukiefferiella sp.	0.40	Goeridae	0.92
Eurycercus	0.33	Gomphidae	2.17
Eurycercus lamellatus	0.33	Gomphidae sp.	1.00
Eusimulium	0.25	*Gomphus*	3.00
Eusimulium costatum	0.25	*Gomphus pulchellus*	3.00
Ferrissia	1.00	*Gomphus* sp.	3.00
Ferrissia wautieri	1.00	*Gomphus vulgatissimus*	3.00
Galba	1.00	*Grammotaulius*	1.00
Galba truncatula	1.00	*Grammotaulius nigropunctatus*	1.00
Gammaridae	0.66	*Grammotaulius nitidus*	1.00
Gammarus	0.73	*Graphoderus*	1.00
Gammarus lacustris	1.00	*Graphoderus austriacus*	1.00
Gammarus pulex	0.75	Grapsidae	1.00
Gammarus roeseli	0.75	*Graptodytes*	1.00
Gammarus salinus	1.00	*Graptodytes flavipes*	1.00
Gammarus sp.	0.75	*Graptodytes granularis*	1.00
Gammarus tigrinus	0.33	*Graptodytes pictus*	1.00
Gammarus zaddachi	0.50	*Graptodytes* sp.	1.00
Gastropoda/Prosobranchia	1.01	*Graptodytes versicolor*	1.00
Gastropoda/Pulmonata	0.94	*Gyraulus*	0.88
Gerridae	0.48	*Gyraulus acronicus*	1.00
Gerridae sp.	0.50	*Gyraulus albus*	1.00
Gerris	0.48	*Gyraulus crista*	0.30
Gerris argentatus	0.50	*Gyraulus laevis*	1.00
Gerris costai	0.50	*Gyraulus parvus*	1.00
Gerris gibbifer	0.50	*Gyraulus* sp.	1.00
Gerris lacustris	0.50	Gyrinidae	1.00
Gerris lateralis	0.50	Gyrinidae sp.	1.00
Gerris najas	0.50	*Gyrinus*	1.00
Gerris odontogaster	0.50	*Gyrinus aeratus*	1.00
Gerris paludum	0.50	*Gyrinus caspius*	1.00
Gerris sp.	0.30	*Gyrinus distinctus*	1.00
Gerris thoracicus	0.50	*Gyrinus marinus*	1.00
Glaenocorisa	0.50	*Gyrinus minutus*	1.00
Glaenocorisa propinqua	0.50	*Gyrinus natator*	1.00
Glossiphonia	0.75	*Gyrinus obsoletus*	1.00
Glossiphonia complanata	1.00	*Gyrinus opacus*	1.00
Glossiphonia heteroclita	0.50	*Gyrinus paykulli*	1.00
Glossiphoniidae	0.79	*Gyrinus* sp.	1.00
Glossiphoniidae sp.	0.50	*Gyrinus substriatus*	1.00
Glossosomatidae	0.63	*Gyrinus urinator*	1.00
Glyphotaelius	1.00	*Habrophlebia*	1.00
Glyphotaelius pellucidus	1.00	*Habrophlebia fusca*	1.00

Habrophlebia lauta	1.00	*Helophorus aequalis*	1.00
Habrophlebia sp.	1.00	*Helophorus alternans*	1.00
Haementeria	1.00	*Helophorus aquaticus*	1.00
Haementeria costata	1.00	*Helophorus arvernicus*	1.00
Haementeria sp.	1.00	*Helophorus asperatus*	1.00
Haemopidae	1.00	*Helophorus brevipalpis*	1.00
Haemopis	1.00	*Helophorus dorsalis*	1.00
Haemopis sanguisuga	1.00	*Helophorus flavipes*	1.00
Halesus	1.00	*Helophorus grandis*	1.00
Halesus digitatus	1.00	*Helophorus granularis*	1.00
Halesus radiatus	1.00	*Helophorus griseus*	1.00
Halesus sp.	1.00	*Helophorus guttulus*	1.00
Halesus tesselatus	1.00	*Helophorus longitarsis*	1.00
Haliplidae	1.12	*Helophorus minutus*	1.00
Haliplidae sp.	1.00	*Helophorus nanus*	1.00
Haliplus	1.07	*Helophorus obscurus*	1.00
Haliplus confinis	1.00	*Helophorus pumilio*	1.00
Haliplus flavicollis	1.00	*Helophorus* sp.	1.00
Haliplus fluviatilis	1.00	*Helophorus strigifrons*	1.00
Haliplus fulvus	1.00	*Hemerodromia*	0.50
Haliplus heydeni	1.00	*Hemerodromia* sp.	0.50
Haliplus immaculatus	1.00	*Hemiclepsis*	1.00
Haliplus laminatus	1.00	*Hemiclepsis marginata*	1.00
Haliplus lineatocollis	2.00	*Heptagenia*	0.80
Haliplus lineolatus	1.00	*Heptagenia flava*	0.50
Haliplus obliquus	1.00	*Heptagenia fuscogrisea*	1.00
Haliplus ruficollis	1.00	*Heptagenia longicauda*	0.50
Haliplus sp.	1.00	*Heptagenia* sp.	1.00
Haliplus variegatus	1.00	*Heptagenia sulphurea*	1.00
Haliplus wehnckei	1.00	Heptageniidae	0.66
Haplotaxida	0.46	*Hesperocorixa*	0.56
Haplotaxidae	0.50	*Hesperocorixa castanea*	0.50
Haplotaxis	0.50	*Hesperocorixa castanea/moesta*	0.50
Haplotaxis gordioides	0.50	*Hesperocorixa linnaei*	0.50
Hebridae	1.00	*Hesperocorixa linnei*	0.50
Hebrus	1.00	*Hesperocorixa moesta*	0.50
Hebrus pusillus	1.00	*Hesperocorixa sahlbergi*	0.50
Hebrus ruficeps	1.00	*Hesperocorixa* sp.	0.90
Helobdella	0.50	Heteroceridae	1.00
Helobdella sp.	0.50	*Heterocerus*	1.00
Helobdella stagnalis	0.50	*Heterocerus fenestratus*	1.00
Helochares	1.00	*Heterocerus* sp.	1.00
Helochares lividus	1.00	*Heteroptera*	0.66
Helochares obscurus	1.00	*Hexatomini*	1.00
Helochares punctatus	1.00	*Hexatomini* sp.	1.00
Helodes	1.00	*Hippeutis*	1.00
Helodes sp.	1.00	*Hippeutis complanatus*	1.00
Helophorus	1.00	*Hirudinea*	0.87

Holocentropus	1.00	*Hydroporus glabriusculus*	1.00
Holocentropus dubius	1.00	*Hydroporus gyllenhali*	1.00
Holocentropus picicornis	1.00	*Hydroporus incognitus*	1.00
Holocentropus sp.	1.00	*Hydroporus latus*	1.00
Holocentropus stagnalis	1.00	*Hydroporus longicornis*	1.00
Hydaticus	1.00	*Hydroporus longulus*	1.00
Hydaticus seminiger	1.00	*Hydroporus marginatus*	1.00
Hydaticus transversalis	1.00	*Hydroporus melanarius*	1.00
Hydatophylax	1.00	*Hydroporus memnonius*	1.00
Hydatophylax infumatus	1.00	*Hydroporus morio*	1.00
Hydracarina	1.00	*Hydroporus nanus*	1.00
Hydracarina sp.	1.00	*Hydroporus neglectus*	1.00
Hydraena	1.00	*Hydroporus nigrita*	1.00
Hydraena britteni	1.00	*Hydroporus obscurus*	1.00
Hydraena gracilis	1.00	*Hydroporus palustris*	1.00
Hydraena melas	1.00	*Hydroporus planus*	1.00
Hydraena nigrita	1.00	*Hydroporus pubescens*	1.00
Hydraena palustris	1.00	*Hydroporus rufifrons*	1.00
Hydraena riparia	1.00	*Hydroporus* sp.	1.00
Hydraena sp.	1.00	*Hydroporus striola*	1.00
Hydraena testacea	1.00	*Hydroporus tesselatus*	1.00
Hydraenidae	1.00	*Hydroporus tristis*	1.00
Hydraenidae sp.	1.00	*Hydroporus umbrosus*	1.00
Hydrobiidae	0.88	*Hydropsyche*	1.00
Hydrobius	1.00	*Hydropsyche angustipennis*	1.00
Hydrobius fuscipes	1.00	*Hydropsyche bulgaromanorum*	1.00
Hydrochara	1.00	*Hydropsyche contubernalis*	1.00
Hydrochara caraboides	1.00	*Hydropsyche incognita*	1.00
Hydrochus	1.00	*Hydropsyche instabilis*	1.00
Hydrochus angustatus	1.00	*Hydropsyche pellucidula*	1.00
Hydrochus carinatus	1.00	*Hydropsyche saxonica*	1.00
Hydrochus elongatus	1.00	*Hydropsyche siltalai*	1.00
Hydrochus nitidicollis	1.00	*Hydropsyche* sp.	1.00
Hydroglyphus	1.00	*Hydropsychidae*	1.00
Hydroglyphus geminus	1.00	*Hydroptila*	0.50
Hydrometra	1.00	*Hydroptila maclachlani*	0.50
Hydrometra stagnorum	1.00	*Hydroptila* sp.	0.50
Hydrometridae	1.00	*Hydroptila sparsa*	0.50
Hydrophilidae	1.00	*Hydroptila vectis*	0.50
Hydrophilidae sp.	1.00	Hydroptilidae	0.65
Hydrophilus	1.00	Hydroptilidae sp.	0.75
Hydrophilus piceus	1.00	*Hydrovatus*	1.00
Hydroporus	1.00	*Hydrovatus clypealis*	1.00
Hydroporus angustatus	1.00	*Hygrobia*	1.00
Hydroporus cantabricus	1.00	*Hygrobia hermanni*	1.00
Hydroporus discretus	1.00	Hygrobiidae	1.00
Hydroporus elongatus	1.00	*Hygrotus*	1.00
Hydroporus erythrocephalus	1.00	*Hygrotus decoratus*	1.00

Hygrotus impressopunctatus	1.00	*Laccobius* sp.	1.00
Hygrotus inaequalis	1.00	*Laccobius striatulus*	1.00
Hygrotus quinquelineatus	1.00	*Laccophilus*	1.00
Hygrotus varius	1.00	*Laccophilus hyalinus*	1.00
Hygrotus versicolor	1.00	*Laccophilus minutus*	1.00
Hypania	0.50	*Laccophilus* sp.	1.00
Hypania invalida	0.50	*Lasiocephala*	1.00
Hyphydrus	1.00	*Lasiocephala basalis*	1.00
Hyphydrus ferrugineus	1.00	Lepidoptera	1.00
Hyphydrus ovatus	1.00	*Lepidostoma*	1.00
Hyphydrus sp.	1.00	*Lepidostoma hirtum*	1.00
Hyporhyacophila	1.00	Lepidostomatidae	1.00
Hyporhyacophila sp.	1.00	Leptoceridae	1.00
Ilybius	1.00	Leptoceridae sp.	1.00
Ilybius aenescens	1.00	*Leptocerus*	1.00
Ilybius angustior	1.00	*Leptocerus lusitanicus*	1.00
Ilybius ater	1.00	*Leptocerus tineiformis*	1.00
Ilybius fenestratus	1.00	*Leptodora*	0.33
Ilybius fuliginosus	1.00	*Leptodora kindtii*	0.33
Ilybius guttiger	1.00	Leptodoridae	0.33
Ilybius obscurus	1.00	*Leptophlebia*	1.00
Ilybius quadriguttatus	1.00	*Leptophlebia marginata*	1.00
Ilybius sp.	1.00	*Leptophlebia vespertina*	1.00
Ilybius subaeneus	1.00	Leptophlebiidae	1.00
Ilyocoris	1.00	*Lestes*	1.00
Ilyocoris cimicoides	1.00	*Lestes dryas*	1.00
Ironoquia	1.00	*Lestes sponsa*	1.00
Ironoquia dubia	1.00	*Lestes virens*	1.00
Ischnura	0.67	*Lestes viridis*	1.00
Ischnura elegans	0.50	Lestidae	1.00
Ischnura pumilio	0.50	*Leucorrhinia*	1.50
Ischnura sp.	1.00	*Leucorrhinia dubia*	1.00
Isoperla	1.00	*Leucorrhinia pectoralis*	2.00
Isoperla grammatica	1.00	*Leuctra*	1.00
Isoperla oxylepis	1.00	*Leuctra digitata*	1.00
Isoperla sp.	1.00	*Leuctra fusca*	1.00
Isopoda	0.46	*Leuctra geniculata*	1.00
Isoptena	1.00	*Leuctra hippopus*	1.00
Isoptena serricornis	1.00	*Leuctra moselyi*	1.00
Ithytrichia	0.50	*Leuctra nigra*	1.00
Ithytrichia lamellaris	0.50	*Leuctra* sp.	1.00
Laccobius	1.00	Leuctridae	1.00
Laccobius alutaceus	1.00	*Libellula*	1.50
Laccobius atratus	1.00	*Libellula depressa*	1.00
Laccobius biguttatus	1.00	*Libellula quadrimaculata*	2.00
Laccobius bipunctatus	1.00	Libellulidae	1.31
Laccobius minutus	1.00	Libellulidae sp.	1.30
Laccobius sinuatus	1.00	*Limnebius*	1.00

Limnebius atomus	1.00	*Limnodrilus claparedeanus*	0.50
Limnebius crinifer	1.00	*Limnodrilus hoffmeisteri*	0.50
Limnebius nitidus	1.00	*Limnodrilus* sp.	0.33
Limnebius papposus	1.00	*Limnophora*	0.50
Limnebius parvulus	1.00	*Limnophora* sp.	0.50
Limnebius sp.	1.00	*Limnoxenus*	1.00
Limnebius truncatellus	1.00	*Limnoxenus niger*	1.00
Limnephilidae	1.00	Limoniidae	1.00
Limnephilidae sp.	1.00	Limoniidae sp.	1.00
Limnephilidae sp. (Junglarven)	1.00	*Lithax*	1.00
Limnephilus	0.94	*Lithax obscurus*	1.00
Limnephilus affinis	1.00	Lumbricidae	0.33
Limnephilus auricula	0.50	*Lumbricina*	0.40
Limnephilus binotatus	1.00	*Lumbriculidae*	0.42
Limnephilus bipunctatus	1.00	*Lumbriculus*	0.33
Limnephilus bipunctatus/centralis	1.00	*Lumbriculus* sp.	0.33
Limnephilus borealis	1.00	*Lumbriculus variegatus*	0.33
Limnephilus centralis	1.00	Lycosidae	1.00
Limnephilus decipens	1.00	*Lymnaea*	1.00
Limnephilus elegans	1.00	*Lymnaea auricularia*	1.00
Limnephilus extricatus	1.00	*Lymnaea glabra*	1.00
Limnephilus flavicornis	0.50	*Lymnaea palustris*	1.00
Limnephilus flavicornis/marmoratus	1.00	*Lymnaea peregra*	1.00
Limnephilus fuscicornis	1.00	*Lymnaea* sp.	1.00
Limnephilus griseus	1.00	*Lymnaea stagnalis*	1.00
Limnephilus hirsutus	1.00	*Lymnaea truncatula*	1.00
Limnephilus ignavus	1.00	Lymnaeidae	0.94
Limnephilus incisus	1.00	*Lype*	0.50
Limnephilus lunatus	1.00	*Lype phaeopa*	0.50
Limnephilus luridus	1.00	*Lype reducta*	0.50
Limnephilus marmoratus	1.00	*Lype* sp.	0.50
Limnephilus nigriceps	1.00	*Lype unicolor*	0.50
Limnephilus politus	1.00	*Malacocerus*	0.50
Limnephilus rhombicus	0.50	*Malacocerus tetracerus*	0.50
Limnephilus sp.	0.90	*Malacostraca*	0.67
Limnephilus sparsus	1.00	*Marenzelleria*	0.50
Limnephilus stigma	1.00	*Marenzelleria viridis*	0.50
Limnephilus subcentralis	1.00	*Marstoniopsis*	1.00
Limnephilus vittatus	1.00	*Marstoniopsis scholtzi*	1.00
Limnichidae	1.00	*Maxillopoda/Branchiura*	0.40
Limnichus	1.00	*Megaloptera*	1.63
Limnichus pygmaeus	1.00	*Megasternum*	1.00
Limnius	2.00	*Megasternum obscurum*	1.00
Limnius sp.	2.00	*Melampophylax*	1.00
Limnius volckmari	2.00	*Melampophylax mucoreus*	1.00
Limnodrilus	0.44	Mesogastropoda	1.01

Mesovelia	1.00	*Nais elinguis*	0.50
Mesovelia furcata	1.00	*Nais* sp.	0.50
Mesoveliidae	1.00	*Nanocladius*	0.40
Metreletus	1.00	*Nanocladius bicolor*	0.40
Metreletus balcanicus	1.00	*Nanocladius* sp.	0.40
Micronecta	0.50	Naucoridae	1.00
Micronecta minutissima	0.50	*Nebrioporus*	1.00
Micronecta poweri	0.50	*Nebrioporus assimilis*	1.00
Micronecta scholtzi	0.50	*Nebrioporus depressus*	1.00
Micronecta sp.	0.50	*Nebrioporus elegans*	1.00
Micropsectra	0.40	*Nebrioporus* sp.	1.00
Micropsectra sp.	0.40	*Nemotaulius*	1.00
Micropterna	1.00	*Nemotaulius punctatolineatus*	1.00
Micropterna lateralis	1.00	*Nemoura*	1.00
Micropterna nycterobia	1.00	*Nemoura avicularis*	1.00
Micropterna sequax	1.00	*Nemoura cambrica*	1.00
Micropterna sp.	1.00	*Nemoura cinerea*	1.00
Microtendipes	1.00	*Nemoura dubitans*	1.00
Microtendipes sp.	1.00	*Nemoura flexuosa*	1.00
Microvelia	0.83	*Nemoura* sp.	1.00
Microvelia buenoi	1.00	Nemouridae	1.00
Microvelia pygmaea	1.00	*Nemurella*	1.00
Microvelia reticulata	0.50	*Nemurella pictetii*	1.00
Mideopsidae	0.50	*Neomysis*	0.30
Mideopsis	0.50	*Neomysis integer*	0.30
Mideopsis orbicularis	0.50	*Nepa*	1.00
Molanna	1.00	*Nepa cinerea*	1.00
Molanna albicans	1.00	*Nepa rubra*	1.00
Molanna angustata	1.00	Nepidae	1.00
Molanna sp.	1.00	Neredidae	0.50
Molannidae	1.00	*Nereis*	0.50
Muscidae	0.50	*Nereis diversicolor*	0.50
Musculium	1.00	*Nereis* sp.	0.50
Musculium lacustre	1.00	Neritidae	1.00
Mysida	0.30	*Neureclipsis*	1.50
Mysidae	0.30	*Neureclipsis bimaculata*	1.50
Mystacides	1.00	*Neuroptera*	0.50
Mystacides azurea	1.00	*Nigrobaetis*	0.33
Mystacides longicornis	1.00	*Nigrobaetis niger*	0.33
Mystacides longicornis/nigra	1.00	Niphargidae	0.42
Mystacides nigra	1.00	*Niphargus*	0.42
Mystacides sp.	1.00	*Niphargus aquilex*	0.50
Myxas	1.00	*Niphargus* sp.	0.33
Myxas glutinosa	1.00	*Normandia*	2.00
Naididae	0.30	*Normandia nitens*	2.00
Naididae sp.	0.30	Noteridae	1.00
Nais	0.44	*Noterus*	1.00
Nais communis	0.33	*Noterus clavicornis*	1.00

Noterus crassicornis	1.00	*Orconectes limosus*	2.00
Notidobia	1.00	*Orectochilus*	1.00
Notidobia ciliaris	1.00	*Orectochilus villosus*	1.00
Notidobia sp.	1.00	*Oreodytes*	1.00
Notonecta	1.00	*Oreodytes sanmarki*	1.00
Notonecta glauca	1.00	*Oreodytes sanmarkii*	1.00
Notonecta maculata	1.00	*Oreodytes septentrionalis*	1.00
Notonecta marmorea	1.00	*Orthetrum*	2.00
Notonecta obliqua	1.00	*Orthetrum brunneum*	2.00
Notonecta sp.	1.00	*Orthetrum cancellatum*	2.00
Notonecta viridis	1.00	*Orthetrum coerulescens*	2.00
Notonectidae	1.00	*Orthocladius*	0.33
Notonectidae sp.	1.00	*Orthocladius* sp.	0.33
Ochthebius	1.00	*Oulimnius*	2.00
Ochthebius bicolon	1.00	*Oulimnius major*	2.00
Ochthebius dilatatus	1.00	*Oulimnius tuberculatus*	2.00
Ochthebius exsculptus	1.00	*Oxyethira*	0.50
Ochthebius marinus	1.00	*Oxyethira falcata*	0.50
Ochthebius minimus	1.00	*Oxyethira* sp.	0.50
Ochthebius nanus	1.00	*Pacifastacus*	1.00
Ochthebius punctatus	1.00	*Pacifastacus leniusculus*	1.00
Ochthebius pusillus	1.00	*Palaemon*	1.00
Ochthebius viridis	1.00	*Palaemon longirostris*	1.00
Odagmia	0.25	*Palaemonetes*	1.00
Odagmia mitidifrons	0.25	*Palaemonetes varians*	1.00
Odagmia ornata	0.25	Palaemonidae	1.00
Odonata	1.46	*Parachironomus*	0.40
Odontoceridae	1.00	*Parachironomus* sp.	0.40
Odontocerum	1.00	*Paracorixa*	0.50
Odontocerum albicorne	1.00	*Paracorixa concinna*	0.50
Oecetis	1.00	*Paracymus*	1.00
Oecetis furva	1.00	*Paracymus scutellaris*	1.00
Oecetis lacustris	1.00	*Paraleptophlebia*	1.00
Oecetis ochracea	1.00	*Paraleptophlebia cincta*	1.00
Oecetis sp.	1.00	*Paraleptophlebia* sp.	1.00
Oecetis testacea	1.00	*Paraleptophlebia submarginata*	1.00
Oligochaeta	0.33	*Paraleptophlebia werneri*	1.00
Oligochaeta sp.	0.33	*Parapoynx*	1.00
Oligoplectrum	1.00	*Parapoynx stratiotata*	1.00
Oligoplectrum maculatum	1.00	*Pararhyacophila*	1.00
Oligostomis	1.00	*Pararhyacophila* sp.	1.00
Oligostomis reticulata	1.00	*Paratanytarsus*	0.40
Oligotricha	1.00	*Paratanytarsus* sp.	0.40
Oligotricha striata	1.00	*Paratendipes*	0.40
Ophiogomphus	1.50	*Paratendipes* sp.	0.40
Ophiogomphus cecilia	2.00	Pediciidae	1.00
Ophiogomphus serpentinus	1.00	*Peloscolex*	0.50
Orconectes	2.00	*Peloscolex* sp.	0.50

Peltodytes	2.00	*Pisidium subtruncatum*	1.00
Peltodytes caesus	2.00	*Pisidium supinum*	1.00
Pentaneurini	0.33	*Planaria*	1.00
Pentaneurini sp.	0.33	*Planaria torva*	1.00
Perla	1.00	Planariidae	0.83
Perla bipunctata	1.00	Planariidae sp.	1.00
Perlidae	1.00	*Planorbarius*	1.00
Perlodes	1.00	*Planorbarius corneus*	1.00
Perlodes microcephala	1.00	Planorbidae	1.00
Perlodes microcephalus	1.00	Planorbidae sp.	1.00
Perlodes sp.	1.00	*Planorbis*	0.83
Perlodidae	1.00	*Planorbis carinatus*	1.00
Phaenopsectra	0.40	*Planorbis planorbis*	0.50
Phaenopsectra sp.	0.40	*Planorbis* sp.	1.00
Phagocata	1.00	*Platambus*	1.00
Phagocata sp.	1.00	*Platambus maculatus*	1.00
Philopotamidae	1.17	Platycnemididae	1.00
Philopotamus	1.00	*Platycnemis*	1.00
Philopotamus sp.	1.00	*Platycnemis pennipes*	1.00
Phryganea	1.00	*Platycnemis* sp.	1.00
Phryganea bipunctata	1.00	*Plea*	1.00
Phryganea grandis	1.00	*Plea leachi*	1.00
Phryganea sp.	1.00	*Plea minutissima*	1.00
Phryganeidae	1.00	*Plecoptera*	1.00
Phryganeidae sp.	1.00	*Plectrocnemia*	1.00
Phyllodocida	0.50	*Plectrocnemia conspersa*	1.00
Physa	1.00	*Plectrocnemia geniculata*	1.00
Physa acuta	1.00	*Plectrocnemia* sp.	1.00
Physa fontinalis	1.00	Pleidae	1.00
Physa heterostropha	1.00	Pleidae sp.	1.00
Physa sp.	1.00	*Poecilobothrus*	1.00
Physella	1.00	*Poecilobothrus* sp.	1.00
Physella acuta	1.00	*Polycelis*	0.67
Physidae	1.00	*Polycelis felina*	1.00
Physidae sp.	1.00	*Polycelis nigra*	1.00
Pirata	1.00	*Polycelis nigra/tenuis*	0.33
Pirata piraticus	1.00	*Polycelis tenuis*	0.33
Piscicola	1.00	*Polycentropodidae*	1.04
Piscicola geometra	1.00	*Polycentropus*	1.00
Piscicolidae	1.00	*Polycentropus flavomaculatus*	1.00
Pisidiidae	0.83	*Polycentropus irroratus*	1.00
Pisidium	0.81	*Polycentropus kingi*	1.00
Pisidium amnicum	1.00	*Polycentropus* sp.	1.00
Pisidium casertanum	0.50	*Polychaeta*	0.50
Pisidium henslowanum	1.00	Polymitarcyidae	1.00
Pisidium nitidum	0.50	*Polypedilum*	0.40
Pisidium pseudosphaerium	0.50	*Polypedilum* sp.	0.40
Pisidium sp.	1.00	*Porhydrus*	1.00

Porhydrus lineatus	1.00	*Psychomyia pusilla*	0.50
Porhydrus sp.	1.00	Psychomyiidae	0.50
Portunidae	1.00	*Ptychoptera*	1.00
Potamanthidae	1.00	*Ptychoptera* sp.	1.00
Potamanthus	1.00	Ptychopteridae	0.75
Potamanthus luteus	1.00	Ptychopteridae sp.	0.50
Potamonectes	1.00	Pyralidae	1.00
Potamonectes assimilis	1.00	Pyralidae sp.	1.00
Potamonectes depressus	1.00	*Pyrrhosoma*	1.00
Potamonectes elegans	1.00	*Pyrrhosoma nymphula*	1.00
Potamonectes griseostriatus	1.00	*Radix*	0.75
Potamophylax	1.00	*Radix auricularia*	1.00
Potamophylax cingulatus	1.00	*Radix ovata*	0.50
Potamophylax latipennis	1.00	*Radix peregra*	0.50
Potamophylax luctuosus	1.00	*Radix* sp.	1.00
Potamophylax nigricornis	1.00	*Ranatra*	1.00
Potamophylax rotundipennis	1.00	*Ranatra linearis*	1.00
Potamophylax sp.	1.00	*Rhagio*	0.50
Potamopyrgus	0.75	*Rhagio* sp.	0.50
Potamopyrgus antipodarum	0.50	Rhagionidae	0.50
Potamopyrgus jenkinsi	1.00	Rhagionidae sp.	0.50
Potamothrix	0.50	*Rhantus*	1.00
Potamothrix hammoniensis	0.50	*Rhantus bistriatus*	1.00
Potamothrix moldaviensis	0.50	*Rhantus exsoletus*	1.00
Pristina	0.50	*Rhantus frontalis*	1.00
Pristina idrensis	0.50	*Rhantus grapii*	1.00
Proasellus	0.66	*Rhantus latitans*	1.00
Proasellus coxalis	0.66	*Rhantus notatus*	1.00
Proasellus meridianus	0.66	*Rhantus* sp.	1.00
Procladius	0.40	*Rhantus suturalis*	1.00
Procladius sp.	0.40	*Rhantus suturellus*	1.00
Procloeon	0.75	*Rheocricotopus*	0.40
Procloeon bifidum	0.50	*Rheocricotopus* sp.	0.40
Procloeon pseudorufulum	1.00	*Rheopelopia*	0.40
Prodiamesa	0.38	*Rheopelopia ornata*	0.40
Prodiamesa olivacea	0.40	*Rheotanytarsus*	0.33
Prodiamesa rufovittata	0.40	*Rheotanytarsus* sp.	0.33
Prodiamesa sp.	0.33	*Rhithrogena*	0.75
Protonemura	1.00	*Rhithrogena semicolorata*	0.50
Protonemura meyeri	1.00	*Rhithrogena* sp.	1.00
Protonemura sp.	1.00	*Rhyacodrilus*	0.50
Psammoryctides	0.50	*Rhyacodrilus coccineus*	0.50
Psammoryctides albicola	0.50	*Rhyacophila*	1.00
Psammoryctides barbatus	0.50	*Rhyacophila dorsalis*	1.00
Psychodidae	0.25	*Rhyacophila fasciata*	1.00
Psychodidae sp.	0.25	*Rhyacophila nubila*	1.00
Psychomyia	0.50	*Rhyacophila obliterata*	1.00
Psychomyia fragilis	0.50	*Rhyacophila praemorsa*	1.00

Rhyacophila sp.	1.00	*Sigara nigrolineata*	0.50
Rhyacophila vulgaris	1.00	*Sigara scotti*	0.50
Rhyacophilidae	1.00	*Sigara semistriata*	0.50
Rhynchelmis	0.50	*Sigara* sp.	1.00
Rhynchelmis limosella	0.50	*Sigara stagnalis*	0.50
Rhynchobdellida	0.81	*Sigara striata*	0.50
Riolus	2.00	*Sigara venusta*	0.50
Riolus cupreus	2.00	*Silo*	0.88
Riolus sp.	2.00	*Silo nigricornis*	0.50
Riolus subviolaceus	2.00	*Silo pallipes*	1.00
Scarodytes	1.00	*Silo piceus*	1.00
Scarodytes halensis	1.00	*Silo* sp.	1.00
Scarodytes sp.	1.00	Simuliidae	0.30
Sciomyzidae	0.50	Simuliidae sp.	0.30
Sciomyzidae sp.	0.50	*Simulium*	0.27
Scirtes	1.00	*Simulium (Boophth.) erythrocephal*	0.25
Scirtes sp.	1.00	*Simulium (Eusimulium) angustipes*	0.27
Scirtidae	1.00	*Simulium (Eusimulium) aureum*	0.25
Scirtidae sp.	1.00	*Simulium (Eusimulium) aureum* Group	0.25
Segmentia	1.00	*Simulium (Nevermannia) costatum*	0.30
Segmentia nitida	1.00	*Simulium (Nevermannia) latigonium*	0.25
Segmentina	1.00	*Simulium (Nevermannia) lundstromi*	0.25
Segmentina nitida	1.00	*Simulium (Nevermannia) vernum*	0.25
Seriata	0.92	*Simulium (Simulium) argyreatum*	0.25
Sericostoma	1.00	*Simulium (Simulium) intermedium*	0.25
Sericostoma flavicorne	1.00	*Simulium (Simulium) morsitans*	0.50
Sericostoma personatum	1.00	*Simulium (Simulium) noelleri*	0.50
Sericostoma sp.	1.00	*Simulium (Simulium) ornatum*	0.25
Sericostomatidae	1.00	*Simulium (Simulium) reptans*	0.25
Sericostomatidae sp.	1.00	*Simulium (Simulium) rostratum*	0.27
Serratella	1.00	*Simulium (Simulium)* sp.	0.25
Serratella ignita	1.00	*Simulium (Simulium) trifasciatum*	0.25
Sialidae	1.63	*Simulium (Wilhelmia) equinum*	0.25
Sialis	1.63	*Simulium (Wilhelmia) lineatum*	0.25
Sialis fuliginosa	2.00	*Simulium brevicaule*	0.25
Sialis lutaria	1.00	*Simulium dunfellense*	0.25
Sialis nigripes	2.00	*Simulium salopiense*	0.25
Sialis sp.	1.50	*Simulium* sp.	0.25
Sigara	0.53	*Simulium spinosum*	0.25
Sigara concinna	0.50	Siphlonuridae	1.00
Sigara distincta	0.50	*Siphlonurus*	1.00
Sigara dorsalis	0.50	*Siphlonurus aestivalis*	1.00
Sigara falleni	0.50	*Siphlonurus armatus*	1.00
Sigara fallenoides	0.50	*Siphlonurus lacustris*	1.00
Sigara fossarum	0.50	*Siphlonurus* sp.	1.00
Sigara hellensi	0.50	*Siphonoperla*	1.00
Sigara lateralis	0.50	*Siphonoperla* sp.	1.00
Sigara limitata	0.50	*Siphonoperla torrentium*	1.00

Sisyra	0.50	*Sympetrum flaveolum*	1.00
Sisyra sp.	0.50	*Sympetrum nigrescens*	1.00
Sisyridae	0.50	*Sympetrum pedemontanum*	1.00
Somatochlora	1.00	*Sympetrum sanguineum*	1.00
Somatochlora metallica	1.00	*Sympetrum* sp.	1.00
Somatochlora sp.	1.00	*Sympetrum striolatum*	1.00
Spercheus	1.00	*Sympetrum vulgatum*	1.00
Spercheus emarginatus	1.00	*Symposiocladius*	0.40
Sperchon	0.50	*Symposiocladius lignicola*	0.40
Sperchon sp.	0.50	Syrphidae	0.25
Sperchonidae	0.50	Syrphidae sp.	0.25
Sphaeriidae	0.75	Tabanidae	0.75
Sphaerium	0.75	Tabanidae sp.	1.00
Sphaerium corneum	0.50	*Tabanus*	0.50
Sphaerium lacustre	0.75	*Tabanus* sp.	0.50
Sphaerium rivicola	1.00	Taeniopterygidae	1.00
Sphaerium sp.	0.75	*Taeniopteryx*	1.00
Spionidae	0.50	*Taeniopteryx nebulosa*	1.00
Stagnicola	1.00	Tanypodinae	0.33
Stagnicola corvus	1.00	Tanypodinae sp.	0.33
Stagnicola corvus/palustris	1.00	*Terebellida*	0.50
Stagnicola palustris	1.00	*Theodoxus*	1.00
Stagnicola sp.	1.00	*Theodoxus fluviatilis*	1.00
Stenochironomus	0.40	*Theromyzon*	0.67
Stenochironomus sp.	0.40	*Theromyzon* sp.	1.00
Stenophylax	1.00	*Theromyzon tessulatum*	0.33
Stenophylax permistus	1.00	*Thoracica*	1.00
Stenophylax sp.	1.00	*Tinodes*	0.50
Stictonectes	1.00	*Tinodes pallidulus*	0.50
Stictonectes lepidus	1.00	*Tinodes* sp.	0.50
Stictotarsus	1.00	*Tinodes unicolor*	0.50
Stictotarsus duodecimpustulatus	1.00	*Tinodes waeneri*	0.50
Stratiomyidae	0.33	*Tipula*	0.50
Stratiomyidae sp.	0.33	*Tipula (Yamatotipula) lateralis*	0.50
Stylaria	0.50	*Tipula (Yamatotipula)* sp.	0.50
Stylaria lacustris	0.50	*Tipula lateralis*	0.50
Stylodrilus	0.50	Tipulidae	0.50
Stylodrilus heringianus	0.50	Tipulidae sp.	0.50
Stylommatophora	1.00	*Torleya*	0.50
Succinea	1.00	*Torleya major*	0.50
Succinea putris	1.00	*Triaenodes*	1.00
Succineidae	1.00	*Triaenodes bicolor*	1.00
Suphrodytes	1.00	*Triaenodes* sp.	1.00
Suphrodytes dorsalis	1.00	*Tricholeiochiton*	0.50
Sympecma	1.00	*Tricholeiochiton fagesii*	0.50
Sympecma fusca	1.00	*Trichoptera*	0.94
Sympetrum	1.00	*Trichostegia*	1.00
Sympetrum danae	1.00	*Trichostegia minor*	1.00

Trocheta	1.00	*Valvata macrostoma*	1.00
Trocheta subviridis	1.00	*Valvata piscinalis*	1.00
Tubifex	0.44	*Valvata pulchella*	1.00
Tubifex costatus	0.50	*Valvatidae*	1.00
Tubifex sp.	0.33	*Velia*	0.75
Tubifex tubifex	0.50	*Velia caprai*	0.50
Tubificidae	0.30	*Velia* sp.	1.00
Tubificidae sp.	0.30	Veliidae	0.50
Turbellaria	0.92	Veliidae sp.	0.50
Turbellaria sp.	1.00	*Veneroida*	0.85
Unio	1.00	Viviparidae	1.00
Unio pictorum	1.00	*Viviparus*	1.00
Unio sp.	1.00	*Viviparus contectus*	1.00
Unio tumidus	1.00	*Viviparus fasciatus*	1.00
Unionidae	1.00	*Viviparus viviparus*	1.00
Unionidae sp.	1.00	*Wormaldia*	1.25
Unionoida	1.00	*Wormaldia occipitalis*	1.00
Valvata	1.00	*Wormaldia* sp.	1.50
Valvata cristata	1.00	Zygoptera	1.05

Appendix E. Variability in exposure–response relationships between micro-/mesocosm experiments performed with the same plant protection product

For a few plant protection products (PPPs) only, more than three micro-/mesocosm experiments have been performed that studied a similar exposure regime. The information available for the organophophorus insecticide chlorpyrifos in particular allows the evaluation of effects of a single-pulse exposure regime (Table E.1).

Table E.1: Effect class concentrations (in µg/L) of the most sensitive measurement endpoint in micro-/mesocosm experiments that studied the impact of single-pulse, repeated-pulse and chronic exposure to the insecticide chlorpyrifos. The effect classes are expressed in terms of nominal concentrations. These nominal concentrations generally were within 20 % of the exposure concentrations on the basis of measurements in the application solutions or in the water column of the test systems.

Exposure regime	Effect class 1	Effect class 2	Effect class 3A	Effect class 4–5	Type of test system	Reference; country
Single pulse (peak)	0.1	0.3	1.0	3.0	Outdoor lentic microcosm	Biever et al., 1994; USA
Single pulse (peak)	–	0.1	–	0.9	Outdoor lentic mesocosms	Van den Brink et al., 1996; Netherlands
Single pulse (peak)	0.1	–	–	1.0	Outdoor lentic mesocosms	Lopez-Mancisidor et al., 2007; Spain
Single pulse (peak)	0.1	–	–	1.0	Outdoor lentic mesocosm	Daam et al., 2008; Thailand
Single pulse (peak)	0.1	–	5[(a)]	–	Outdoor lotic mesocosm	Pusey et al., 1994; Australia
Single pulse (peak)	–	–	0.5	6.3	Outdoor lentic mesocosm	Siefert et al., 1989; USA
Single pulse (peak)	0.1	–	1.0	10	Indoor lentic cosm; 16 °C, mesotrophic	Van Wijngaarden et al., 2005; Netherlands
Single pulse (peak)	0.1	–	1.0	–	Indoor lentic cosm; 26 °C, mesotrophic	Van Wijngaarden et al., 2005; Netherlands
Single pulse (peak)	0.1	–	–	1.0	Indoor lentic cosm; 26 °C, eutrophic	Van Wijngaarden et al., 2005; Netherlands
Repeated pulse (4×)	0.033		0.1	1	Outdoor lentic mesocosms	Lopez-Mancisidor et al., 2008; Spain
Constant chronic (28 days)	–	–	–	0.1	Indoor lentic microcosm	Van den Brink et al., 1995; Netherlands
Constant chronic (28 days)	–	0.01[b]	–	0.1[(b)]	Indoor lentic microcosm	Van den Brink et al., 2002; Netherlands

(a): Recovery is fast because of constant input of propagules in experimental stream after pulse exposure.
(b): Exposure to a mixture of chlorpyrifos and lindane; all treatment-related effects were assigned to chlorpyrifos.

For eight aquatic micro-/mesocosm experiments, performed in different parts of the world and/or under different experimental conditions, an effect class 1–2 response was observed at a peak concentration of 0.1 µg chlorpyrifos/L. Note that this is partly due to the fact that similar exposure concentrations were selected by the different experimenters. Nevertheless, the similarity between effect class 1–2 responses between different studies can be explained by the fact that both crustaceans and insects are sensitive to this insecticide and that the communities of the micro-/mesocosm test systems used all contained a reasonably high diversity of these arthropods.

It appears that differences in these effect class 3A concentrations between studies are relatively large. Note, however, that from a regulatory point of view it is fair to make a distinction in recovery of sensitive arthropods between hydrologically isolated test systems (lentic micro-/mesocosms: effect class 3A concentrations ≤ 1.0 µg/L) and the outdoor stream in which a more or less constant inflow of sensitive stream invertebrates was possible (resulting in an effect class 3A concentration of 5 µg/L). It also appears from the chlorpyrifos data presented in Table E.1 that the threshold concentration (effect class 1) of the repeated (4 ×) pulse exposure study is a factor of approximately 3 lower than that of the single exposure studies. Treatment-related effects due to a constant chronic exposure probably occur at concentrations ≥ 0.01 µg chlorpyrifos/L.

For the pyrethroid insecticide lambda-cyhalothrin the majority of micro-/mesocosm experiments available concern repeated application studies (Table E.2). It appears that the variability in effect class 1 ($n = 2$) and effect class 2 ($n = 4$) responses between different studies is remarkably low, while that for effect class 3A ($n = 6$) responses is somewhat higher.

Table E.2: Effect class concentrations (in ng/L) of the most sensitive measurement endpoint in micro/mesocosm experiments that studied the impact of pulsed exposures to the insecticide lambda-cyhalothrin. The effect classes are expressed in terms of nominal peak concentrations. In most studies the nominal concentrations were in accordance with measurements of the test substance in the application solutions.

Exposure regime	Effect class 1	Effect class 2	Effect class 3A	Effect class 4–5	Type of test system	Reference; country
Single pulse	–	–	50	–	Outdoor lotic mesocosms	Heckmann and Friberg, 2005; Denmark
Repeated pulse (12×)	2.7[(a)]	–	–	27[(a)]	Outdoor lentic mesocosms	Hill et al., 1994; USA
Repeated pulse (2×)	4.0[(b)]	–	16[(b)]	85[(b)]	Outdoor lentic mesocosms	Arts et al., 2006; Netherlands
Repeated pulse (5×)	–	10[(b)]	25[(b)]	50[(b)]	Indoor lentic microcosms	Van Wijngaarden et al., 2004; Netherlands
Repeated pulse (3×)	–	10	10	25	Outdoor lentic microcosm	Roessink et al., 2005; Netherlands
Repeated pulse (3×)	–	10	50	–	Outdoor lentic microcosm	Roessink et al., 2005; Netherlands
Repeated pulse (3×)	–	10	10–25	50	Outdoor lentic microcosms	Van Wijngaarden et al., 2006; Netherlands
Repeated pulse (3×)	–	–	–	17	Outdoor lentic mesocosms	Farmer et al., 1995; UK

(a): Experiment was characterised by both spray drift (nominal 1.6 and 16 µg/L) and run-off applications (nominal 4.7 and 47 µg/L). As exposure concentration the median value for the spray drift and run-off application was used.
(b): Exposure to a realistic package of different PPPs used in a specific crop including lambda-cyhalothrin; all treatment-related effects were assigned to lambda-cyhalothrin.

For the pyrethroid insecticide esfenvalerate the majority of micro-/mesocosm experiments available concern pulsed exposures (single or repeated applications) (Table A.3). It appears that the variability in effect class 1–2 ($n = 2$) concentrations of the two repeatedly exposed complex mesocosm studies (including many insect populations) are similar (0.01 µg/L). In addition, an effect class 3A concentration of 0.03 µg/L was observed for the most sensitive endpoint reported (in the paper) in a simple plankton-dominated microcosm test (NSH-NH microcosms of Stampfli et al. (2011). Note that an additional stressor in the form of harvesting a considerable proportion of zooplankton (NSH-H microcosm) resulted in an effect class 2 concentration of 0.03 µg esfenvalerate/L (on the basis of the most sensitive endpoint reported), while the additional stressor in the form of shading (SH-NH microcosms) resulted in a similar an effect class 3A lowest observed effect concentration (LOEC) of 0.03 µg/L esfenvalerate.

Table E.3: Effect class concentrations (in µg/L) of the most sensitive measurement endpoint in micro/mesocosm experiments that studied the impact of (short-term) pulsed exposures to the insecticide esfenvalerate. The effect classes are expressed in terms of nominal peak concentrations. In most studies the nominal concentrations were in accordance with measurements of the test substance in the application solutions

Exposure regime	Effect class 1	Effect class 2	Effect class 3A	Effect class 4–5	Type of test system	Reference; country
Repeated pulse (2×) **esfenvalerate**	–	0.01	–	0.08	Littoral enclosures	Lozano et al., 1992; USA
Repeated pulse (10×) **esfenvalerate**	0.01	–	–	0.25	Outdoor mesocosms	Webber et al., 1992 USA
Singe pulse **esfenvalerate**	–	–	0.03[(a)]	0.3[(a)]	NSH-NH microcosms	Stampfli et al., 2011; Germany
Single pulse **esfenvalerate**	–	0.03[(a),(b)]	–	3[(a),(b)]	NSH-H microcosms	Stampfli et al., 2011; Germany
Single pulse **esfenvalerate**		–	0.03[(a),(b)]	3[(a),(b)]	SH-NH microcosms	Stampfli et al., 2011; Germany

(a): Focus on plankton organisms only.
(b): Additional stressor.

There appears to be limited information on PPP-treated model ecosystems comparing effect class 1 or class 2 concentrations for direct toxic effects as a result of more or less constant chronic exposure. The limited micro-/mesocosm information available for the persistent fungicide carbendazim suggests little variation in effect class 1 concentrations between experiments as a result of a long-term chronic exposure regime (Table E.4).

Table E.4: Effect class concentrations (in µg/L) of the most sensitive measurement endpoint in micro-/mesocosm experiments (fish not present) that studied the impact of more or less constant exposure to the fungicide carbendazim.

Exposure regime	Effect class 1	Effect class 2	Effect class 3	Effect class 4	Type of test system	Reference; country
Long-term	2.6	–	–	26.4	Outdoor microcosms	Daam et al., 2009; Thailand
Long-term	2.2	–	–	20.7	Outdoor mesocosms	Slijkerman et al., 2004; Netherlands
Long-term	3.3	–	–	33.0	Indoor microcosms	Cuppen et al., 2000; Van den Brink et al., 2000; Netherlands

A large dataset is available only for the persistent herbicide atrazine (Table E.5).

Table E.5: Effect class concentrations (in µg/L) of the most sensitive measurement endpoint in micro-/mesocosm experiments that studied the impact of more or less constant long-term exposure to the herbicide atrazine.

Exposure regime	Effect class 1	Effect class 2	Effect class 3B	Effect class 4–5	Type of test system	Reference
Long-term	–	2	–	30	Outdoor lentic mesocosms	Seguin et al., 2001
Long-term	5	–	–	–	Indoor lentic microcosms	Van den Brink et al., 1995
Long-term	–	–	–	–	Indoor lotic microcosms	Gruessner and Watzin, 1996
Long-term	5	–	–	22	Outdoor lentic microcosms	Jüttner et al., 1995
Long-term	–	10	–	100	Indoor lentic microcosm	Johnson 1986
Long-term	5	–	50	100	Indoor lentic microcosm	Brockway et al., 1984
Long-term	10	–	–	32	Indoor lentic microcosms	Pratt et al., 1988
Long-term	–	–	–	10	Indoor lotic microcosms	Kosinski 1984, Kosinski and Merkle 1984
Long-term	14	25	–	80	Indoor lotic microcosms	Nyström et al., 2000
Long-term	–	–	–	14	Indoor lotic microcosms	Muňos et al., 2001
Long-term	–	–	–	15	Experimental swamp	Detenbeck et al., 1996
Long-term	–	–	–	20	Outdoor lentic mesocosms	DeNoyelles et al., 1994 (and literature cited)
Long-term	–	20	–	100	Indoor lentic microcosms	Stay et al., 1989
Long-term	–	–	–	24	Indoor lotic microcosms	Krieger et al., 1988

Long-term	–	–	–	50	Outdoor lentic mesocosms	Fairchild et al., 1994

Data available for atrazine suggest a larger variability in class 1 and class 2 effect concentrations between chronic exposure experiments; however, a larger number of studies is also available. Effect class 1 concentrations could be derived from five different atrazine studies, and effect class 2 concentrations from six studies (Table E.5).

The relatively high variability in effect class 1–2 concentrations for chronic studies with atrazine when compared with those with pulsed exposures to chlorpyrifos and lambda-cyhalothrin might be explained by differences in the toxic mode of action of these substances. Atrazine is a photosystem II inhibitor. According to Guasch and Sabater (1998), inhibition of photosynthesis by atrazine is influenced by ambient light conditions, which most probably considerably varied between the different micro-/mesocosm studies reported in Table E.5. Consequently, the question at stake is whether the results from the chronic micro-/mesocosm studies with atrazine are representative for PPPs with another toxic mode of action.

Appendix F. Minimal detectable difference (MDD)

The data from micro-/mesocosm experiments can be analysed for significance by using the minimal detectable differences (MDD) concept (section 9.3.2.5). This approach can be applied to a given endpoint, at a given time, usually to detect treatment-related changes in abundance/biomass of a sensitive population.

The MDD can be calculated based on the following formula:

Equation 1: $$MDD = \overline{x_1} - \overline{x_2} = t \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Equation 2: $$\%MDD = \frac{MDD}{\overline{x_1}} \cdot 100$$

$\overline{x_1}, \overline{x_2}$ arithmetic mean of controls and treatments
s^2_1, s^2_2 variation of variables for control and treatment
n_1, n_2 number of replicates for control and treatment
t tabulated t-value for t-test

(adapted from Lee and Gurland, 1975)

As an example, the variance of the data (given as coefficient of variation of the control, CV) is calculated for given values of MDD (%) (i.e. 30, 50, 70 and 90 %), assuming that the variation in the control and in the treatment is the same, for different numbers of replicates.

The results are listed in the following table.

Table F.1: Examples of calculations of CV (assumed to have the same value for both control and treatment datasets) calculated for given values of MDD, in various experimental studies with different numbers of replicates (calculations based on two-sided test (α = 0.05)).

Experimental study	n_1 (control)	n_2 (treatment)	MDD (% effect difference)[(a)]	CV (%)
Study 1	7	4	90	78
			70	61
			50	44
			30	26
Study 2	5	3	90	63
			70	49
			50	35
			30	21
Study 3	4	3	90	58
			70	45
			50	32
			30	19
Study 4	4	2	90	49
			70	38
			50	27
			30	16

(a): See Table 31 in GD, Proposal on classes of MDD: 10–50 % effect difference = small effects (class IV); up to 70 % effect difference = medium effects (class III); up to 90 % effect difference = large to medium effects (class II).

Micro-/mesocosm experiments are mostly performed using low numbers of replicates (i.e. four replicates for the control and two to three replicates for the treatment).

The results indicate that in such an experiment (e.g. with four replicates for the control and two to three replicates per treatment), the CV between replicates for both control and treatment should not exceed 32 %, so that a 50 % effect difference between these two datasets is detected.

Appendix G. Worked examples for qualitative uncertainty evaluation

As outlined in chapter 12 on uncertainty, every refined assessment should contain at least a qualitative evaluation of uncertainties. Below in Tables G.1, G.2 and G.3 three examples are presented of how such a qualitative uncertainty analysis may be performed and presented.

Table G.1: Worked example of a qualitative assessment of the uncertainty in tier 1 fish acute (96 h LC_{50} with rainbow trout). Scale for factor of potential under/overestimation of real risk: *< 5; **5–10; ***> 10

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
Intra-laboratory variation	*	Variation in environmental variables within 'test laboratories' through time/space: variations among different testing labs (noise, light, temperature or variability in these parameters); time of year of exposure study; changes in operators (different levels of disturbance)	*	Variation in environmental variables within 'test laboratories' through time/space: variations among different testing labs (noise, light, temperature or variability in these parameters); time of year of exposure study; changes in operators (different levels of disturbance)
Inter-laboratory variation	*/**	Differences in exposure systems (materials, flow rates, vibration, sound pollution, other sources of disturbance); water quality variables (hardness, other aspects of ionic composition e.g. iodine, TDS); feeding protocols and food quality, nutrient/trace element/contaminant composition; operator activity/competence (e.g. tank cleaning, sampling for analytical chemistry)	*/**	Differences in exposure systems (materials, flow rates, vibration, sound pollution, other sources of disturbance); water quality variables (hardness, other aspects of ionic composition e.g. iodine, TDS); feeding protocols and food quality, nutrient/trace element/contaminant composition; operator activity/competence (e.g. tank cleaning, sampling for analytical chemistry)
Intra-species variation	**	Variability among test organisms from different supplier, owing to environmental and genetic differences. E.g. lab test fish are likely to be better fed and unlikely to carry any disease burden, they may be more tolerant of a toxic chemical insult	*	Variability among test organisms from different supplier, owing to environmental and genetic differences

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
Life-stage sensitivity	*	Test is conducted with juvenile rainbow trout which may be less sensitive than earlier life stages though these are covered in chronic testing (i.e. fish ELS)	*	Juvenile rainbow trout (typically 5 cm length) have higher surface area/volume ratio than more mature fish (e.g. reproducing adults) and thus may be subject to higher tissue concentrations during the test duration than larger fish would be
Differences in life-history characteristics of test organism vs. non-target species influencing exposure (e.g. food web position, ecological niche)	**	The acute toxicity test only provides for exposure to the freely dissolved fraction, not via diet. Differences in life-history characteristics between test organism (a fish) and wild species (fish and other aquatic vertebrates e.g. amphibian, adults or larvae) lead to underestimation of risk. For example, surface feeding amphibian larvae may be exposed to higher concentrations for compounds that stratify in the water column.	**	The acute toxicity test only provides for exposure to the freely dissolved fraction, not via diet. Differences in life-history characteristics between test organism (a fish) and wild species (fish and other aquatic vertebrates e.g. amphibian, adults or larvae) that influence exposure may lead to overestimation of the risk. For example species may be more of less able to avoid surface waters contaminated by spray drift (e.g. adult amphibians)
Inter-species variation in toxicant sensitivity	*	Differences in toxicokinetics (ADME); toxicodynamics (enzyme/receptor subtype/affinity). As rainbow trout are typically among the most sensitive fish species, it is more likely that use of Rt will overestimate the risk	***	Differences in toxicokinetics (ADME); toxicodynamics (enzyme/receptor subtype/affinity). As rainbow trout are typically among the most sensitive fish species, it is more likely that use of Rt will overestimate the risk
Overall assessment	The exposure scenario in an acute lab test likely represents a worst case scenario for short-term exposure of fish and other aquatic vertebrates to the freely dissolved fraction of the a.s./PPP in surface waters. While there are a number of significant uncertainties, few are likely to only operate in the direction of underestimate of true risk.			

Table G.2: Worked example of a qualitative assessment of the uncertainty in tier 2B chronic SSD. Scale for factor of potential under-/overestimation of real risk: *< 5; **5–10; ***> 10

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
Intra-laboratory variation	*	Variation in environmental variables within 'test laboratories' through time/space: variations among different testing labs (noise, light, temperature or variability in these parameters); time of year of exposure study; changes in operators (different levels of disturbance)	*	Variation in environmental variables within 'test laboratories' through time/space: variations among different testing labs (noise, light, temperature or variability in these parameters); time of year of exposure study; changes in operators (different levels of disturbance)
Inter-laboratory variation	*/**	Differences in exposure systems (materials, flow rates, vibration, sound pollution, other sources of disturbance); water quality variables (hardness, other aspects of ionic composition, e.g. iodine, TDS); feeding protocols and food quality, nutrient/trace element/contaminant composition; operator activity/competence (e.g. tank cleaning, sampling for analytical chemistry)	*/**	Differences in exposure systems (materials, flow rates, vibration, sound pollution, other sources of disturbance); water quality variables (hardness, other aspects of ionic composition, e.g. iodine, TDS); feeding protocols and food quality, nutrient/trace element/contaminant composition; operator activity/competence (e.g. tank cleaning, sampling for analytical chemistry)
Intra-species variation	**	Variability between lab-bred and externally supplied stocks, and among external suppliers of test organisms, owing to environmental and genetic differences, e.g. strain; variation between lab test organisms and wild organisms, e.g. lab test fish are likely to be better fed and unlikely to carry any disease burden, and they may be more tolerant of a toxic chemical insult	*	Variability among test organisms from different suppliers owing to environmental and genetic differences
Life stage sensitivity and size	*	Variability among contributing data with respect to life stage and size of organism tested, i.e. adult, juvenile, larvae. For example, an SSD combining fish and amphibian data might combine adult/juvenile data from fish but only larval data for amphibians, being the most relevant life stage for edge-of-field surface waters and most freely available. Use of smaller life stages will also influence uptake via surface area/volume ratio	*	Variability among contributing data with respect to life stage and size of organism tested, i.e. adult, juvenile, larvae. For example, an SSD combining fish and amphibian data might combine adult/juvenile data from fish but only larval data for amphibians, being the most relevant life stage for edge-of-field surface waters and most freely available. Use of smaller life stages will also influence uptake via surface area/volume ratio

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
Laboratory exposure scenario	**	The laboratory exposure scenario will most often be via immersion, i.e. exposure to the freely dissolved fraction, and this may result in underestimation of toxicity owing to additional uptake routes	*	The laboratory exposure scenario will represent a worst-case scenario for exposure to the freely, dissolved fraction, i.e. use of flow-through exposure will provide a constant exposure which is unlikely to be representative of more variable exposure in real world edge-of-field surface waters
Exposure duration	**	Variation in test duration (relative to life span of the test organism) may result in under estimation of latency of effects	**	Variation in test duration (relative to life span of the test organism) may result in over estimation of latency of effects
Endpoint comparability	***	Chronic endpoints tend to be developed and validated relative to a given test organism/species, and therefore combining chronic toxicity values from multiple species risks the assumption that endpoints are directly comparable with regard to toxicant sensitivity for multiple species/taxonomic groups in spite of variability in life spans, reproductive mode, behaviour which may be significant. Unlikely to affect results in one direction more than the other	***	Chronic endpoints tend to be developed and validated relative to a given test organism/species, and therefore combining chronic toxicity values from multiple species risks the assumption that endpoints are directly comparable with regard to toxicant sensitivity for multiple species/taxonomic groups in spite of variability in life spans, reproductive mode, behaviour which may be significant. Unlikely to affect results in one direction more than the other
Species selection	*	For pesticides with specific and well-known modes of action, selection of potentially sensitive species may be fairly easy and reliable, while for pesticides with 'biocidal' properties this process may be less clear-cut	*	For pesticides with specific and well-known modes of action, selection of potentially sensitive species may be fairly easy and reliable, while for pesticides with 'biocidal' properties this process may be less clear-cut
Compilation of species data	**/***	Approach to combining multiple toxicity values for a single species (species mean acute value; SMAV) or genus (genus mean acute value; GMAV)—Geomean approach; whether GMAVs are weighted according to how many SMAVs they represent. Unlikely to affect results in one direction more than the other	**/***	Approach to combining multiple toxicity values for a single species (species mean acute value; SMAV) or genus (genus mean acute value; GMAV)—Geomean approach; whether GMAVs are weighted according to how many SMAVs they represent. Unlikely to affect results in one direction more than the other
Censoring of data	**/***	Use of uncensored values, i.e. within the range of other values or only outside of the range; use in goodness of fit testing. Unlikely to affect results in one direction more than the other	**/***	Use of uncensored values, i.e. within the range of other values or only outside of the range; use in goodness of fit testing. Unlikely to affect results in one direction more than the other

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
Curve fitting	**	Assumption of log-normal distribution: whether the log-normal distribution fits the data adequately, whether the curve fitting is optimised for a good fit at the lower end of toxicity values	**	Assumption of log-normal distribution: whether the log-normal distribution fits the data adequately, whether the curve fitting is optimised for a good fit at the lower end of toxicity values
Confidence parameter	***	Use of median HC_5 or lower limit of 95 % confidence interval of the HC_5. With few species and/or sub-optimal fit of the log-normal (or other) distribution, confidence intervals will expand at the ends of the distribution and this will drive the lower limit of the HC_5 confidence band to the left	*	Use of median HC_5 or lower limit of 95 % confidence interval of the HC_5. If the curve fit is not good at the left hand end of the distribution, then use of the median HC_5 and may result in a higher SSD-RAC than would be generated using the $LLHC_5$
Calibration of HC_5 against higher tiers, i.e. experimental ecosystem data	*/**/***	For algae and invertebrates it is possible to calibrate SSD data against micro-/mesocosm data for well-studied compounds, though fewer data are available for doing this with compounds with a novel mode of action. Calibration of SSDs for vascular plants against appropriate micro/mesocosms has not yet been performed. For fish SSDs, it is not possible to calibrate against a surrogate reference tier as micro-/mesocosms typically do not include fish species	*/**/***	For algae and invertebrates it is possible to calibrate SSD data against micro/mesocosm data for well-studied compounds, though fewer data are available for doing this with compounds with a novel mode of action. Calibration of SSDs for vascular plants against appropriate micro-/mesocosms has not yet been performed. For fish SSDs, it is not possible to calibrate against a surrogate reference tier as micro-/mesocosms typically do not include fish species
Overall assessment	Use of SSDs provides more information on the variation of sensitivity among relevant species, and thus offers the chance to reduce the uncertainty in extrapolating from lab tests to the field that could be attributed to interspecies variation in toxicity. While the use of acute data in SSDs is quite well established and validated (e.g. relative to higher tiers), there are fewer examples of SSDs using chronic endpoints, little calibration relative to other tiers, and guidance may be required on the degree of comparability of endpoints combined in a given SSD.			

HC_5, hazardous concentration for 5 % of the species of an SSD; SSD, species sensitivity distribution

Table G.3: Worked example of a qualitative assessment of the uncertainty in tier 3 mesocosm study. Scale for factor of potential under-/overestimation of real risk: *< 5; **5–10; ***> 10

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
Exposure regime tested in micro-/mesocosms	*	Tested exposure regime informed by predicted exposure profiles (FOCUS- or MS-specific scenarios). Predicted exposures by FOCUS/MS specific tools may be biased	*	Worst-case exposure regimes can be studied if known. Effects usually assigned to freely dissolved fraction of chemical (while also other exposure routes may cause the effects, e.g. via sediment exposure)
Enclosed community in micro-/mesocosms	**	Test systems contain only one of the possible communities for edge-of-field surface waters. Species with complex life cycles are often under-represented. Requirement that at least eight representatives of potential sensitive groups should be sufficiently present in sufficient densities to allow NOEC calculations	*	Populations of sensitive species may be introduced that normally may not be common in the edge-of-field aquatic systems at risk. Indirect effects due to species interaction can be studied
Power of test to detect treatment-related effects	*	Due to financial constraints usually a limited number of test systems is used. Variability is an intrinsic property of ecosystems and in real edge-of-field ecosystems variability of relevant endpoints may be high as well	*	Experimental design and availability of control test systems allow exclusion of confounding factors. Power of test can be increased by more efficient sampling techniques and by increasing the number of replicates
Sampling of measurement endpoints	*	Different types of organisms require different techniques which may be labour intensive. Internationally accepted guidance on sampling methods available	*	Different types of organisms require different techniques which is labour intensive. Internationally accepted guidance on sampling methods available
Statistical evaluation	*	The analysis of complex datasets requires specialists. Internationally accepted guidance how to analyse micro-/mesocosm datasets is available	*	The analysis of complex datasets requires specialists. Internationally accepted guidance how to analyse micro-/mesocosm datasets is available
Detecting of threshold levels for sensitive populations	*	Within a taxonomic group the overall difference in toxicological sensitivity between multivoltine and uni-/semivoltine organisms is small. Representatives of the relevant taxonomic groups should be available	*	Latency of effects can be studies due to medium- to long-term observation
Detection of recovery potential	**	Over-representation of short-cycle populations	*	Long-term observations theoretically possible. The test systems can be designed to contain uni-/semivoltine populations of sensitive taxonomic groups
Interpretation of ecological effect data	**	Specialists required due to complex nature of population- and community-level dataset.	**	Specialists required due to complex nature of population- and community-level dataset.

Source of uncertainty	Potential to underestimate the real risk	Explanation	Potential to overestimate the real risk	Explanation
		Internationally accepted guidance available		Internationally accepted guidance available
Overall assessment	Evaluation criteria for a proper micro-/mesocosm test such as that recommended in the Aquatic Guidance Document, as well as the AF proposed to extrapolate results of micro-/mesocosm data to the field, suggest that the true risk for edge-of-field surface waters is fairly well assessed			

AF, assessment factor; MS, Member State; NOEC, no observed effect concentration.

Appendix H. Case studies

Case studies were performed based on three imaginary active substances with realistic properties (an insecticide, a herbicide and a fungicide). The active substances used for the case studies were designed to demonstrate how to apply the most relevant parts of the guidance, especially the higher tier options. Therefore, the case studies show that the presented compounds would not always pass the risk assessment. As the main aim of the case studies was to show the potential refinements of the effect assessments, not all options for refinement that would be possible in terms of exposure were tested and shown, as this is outside the scope of this guidance. Physicochemical properties that were used for performing the step1, 2 and 3 exposure calculations are shown in Table H.1.

The case studies address especially the following:

H.1 Herbicide case study: effect assessment tier 1 to tier 2B, metabolite effect assessment.

H.2 Fungicide case study: effect assessment tier 1 to tier 3.

H.3 Insecticide case study: effect assessment tier 1 to tier 3, details on mesocosm evaluation including minimum detectable difference (MDD) calculations.

Table H.1: Physicochemical properties of model active substances used in the case studies

Property	**Herbicide**	**Fungicide**	**Insecticide**
Crop	Spring cereals	Winter cereals	Apples
Number of applications	1	1	2
Application rate (kg/ha)	0.02	0.75	0.07/0.105
Time between applications (d) (step 2)	–	–	30
Season of application (step 2)	Spring	Spring	Spring and summer
Crop growth stage(s) at application date	BBCH 32–37	BBCH 32	BBCH 10/BBCH 69–71
Molar mass (g/mol)	400	225	250
Water solubility (mg/L) at 20 °C	3000	13	600
Saturated vapour pressure (mPa) at 20 °C	1e-7	0.5	1e-5
DegT50 at 20 °C, pF = 2 in top soil (d)	20	50	100
DegT50 in water (d) at 20 °C	150	10	5
DegT50 in sediment (d) at 20 °C	100	20	100
K_{OC} (L/kg) for soil	40	1700	170
K_{OM} (L/kg) for soil	23	1000	100

H.1 Herbicide M_A case study

H.1.1 M_A (amino acid biosynthesis inhibitor): general information

Herbicide M_A affects the synthesis of essential amino acids in plants and is used as a selective pre- and post-emergence herbicide. It inhibits cell division in roots and shoots of the plant by inhibiting the enzyme acetolactate synthase (ALS), which is a key enzyme in the biosynthesis of branched amino acids. In efficacy trials terrestrial dicotyledonous plants were more sensitive than monocotyledonous plants. In aquatic toxicity tests with macrophytes it was the formation of new shoots and roots that was inhibited by herbicide M_A, while existing roots and shoots hardly suffered any toxic effects.

In terrestrial soils one relevant metabolite met-M_A is formed that also enters surface waters and for which an effect and risk assessment has to be performed as well. First the effect/risk assessment will be performed for herbicide M_A, followed by that for its main relevant metabolite met-M_A.

In this case the effect/risk assessment is performed for the use of herbicide M_A in spring cereals according to the use pattern described in Table H.1. Some typical predicted exposure profiles for edge-of-field surface waters on the basis of FOCUS scenarios and models are presented in Figure H.1.

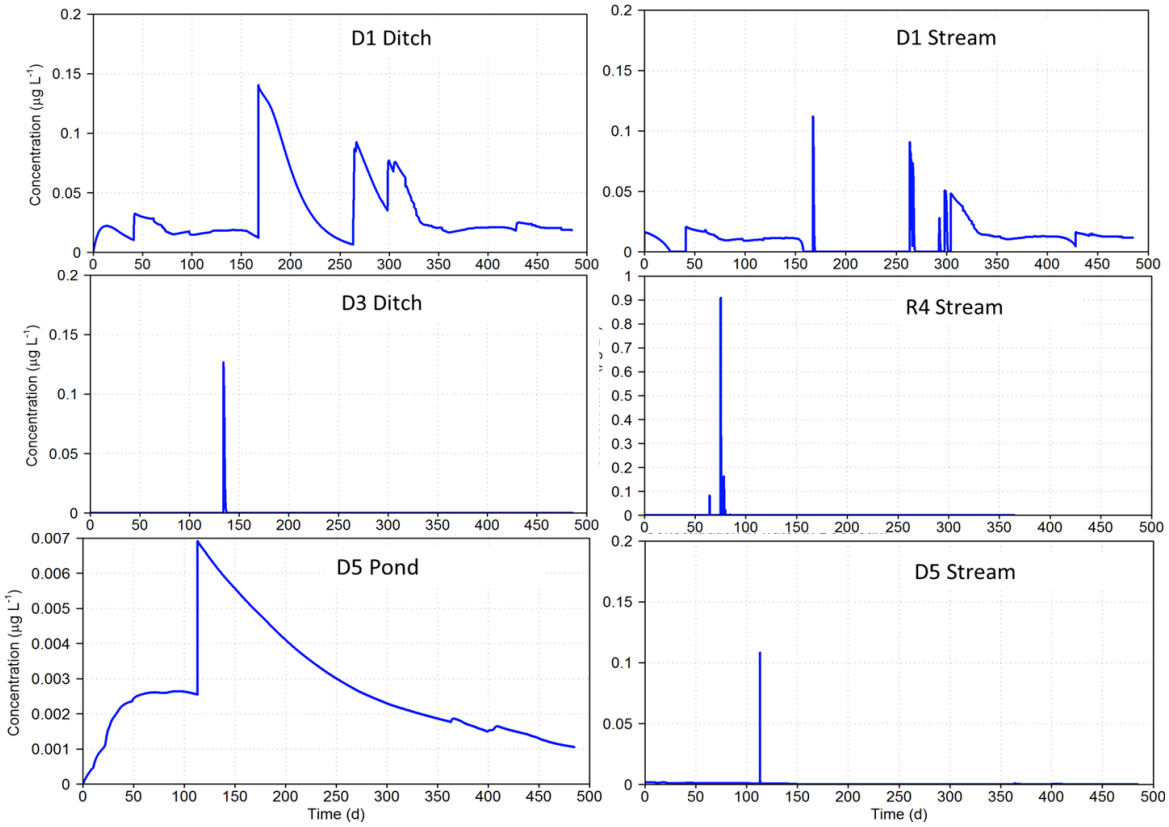


Figure H.1: Typical exposure profiles predicted for herbicide M_A in edge-of-field surface waters on the basis of FOCUS scenarios and models (step 3 calculations)

FOCUS scenario step 3 calculations result in $PEC_{sw;max}$ values in the range of 0.006 (D5 Pond) to 0.569 µg/L (R4 Stream). Considering 90 % drift-reducing nozzles (step 4 calculations) the range does not change as the maximum concentrations are driven by run-off entries.

H.1.2 Herbicide MA effect assessment

H.1.2.1 Tier 1 effect assessment

Toxicity data for standard test species are presented in Table H.2.

Table H.2: Acute and chronic toxicity data for standard test species and herbicide M_A

	Acute $L(E)_{50}$ (µg a.s./L)	**Chronic EC_x/NOEC (µg a.s./L)**	**Endpoint**
Pseudokirchneriella subcapitata	–	119 (72 h E_rC_{50})	Growth rate
Navicula peliculosa	–	95 (96 h E_rC_{50}))	Growth rate
Lemna minor	–	0.718 (7 d E_rC_{50})	Growth rate (frond numbers)
Myriophyllum spicatum	–	0.280 (14 d E_rC_{50})	Growth rate (dry weight of newly formed tissues)
Daphnia magna	> 150 000 (48 h)	150 000 (21 d NOEC)	Mortality
Oncorhynchus mykiss	100 000 (96 h)	60 000 (21 d NOEC)	Mortality

Macrophytes (*Lemna*, *Myriophyllum*) comprise the most sensitive taxonomic group. By applying an AF of 10 to the lowest chronic toxicity value of primary producers and an assessment factor (AF) of 100 to the lowest acute toxicity value for animals, the **tier 1 acute RAC becomes 0.028 µg/L** (application of an AF of 10 to the E_rC_{50} of *Myriophyllum spicatum*).

Comparing this tier 1 RAC with step 3 $PEC_{sw;max}$ values (Table H.3) shows that potential risks may be high for all scenarios, except D4 Pond and D5 Pond. A refined risk assessment is necessary.

Table H.3: Step 3 $PEC_{sw;max}$ values for herbicide M_A in different FOCUS scenarios. The red $PEC_{sw;max}$ values indicate that they are higher than the tier 1 RAC_{sw} (0.028 µg/L) and that under these circumstances the estimated risks are unacceptable. The black and bold values indicate that the $PEC_{sw;max}$ is lower than the RAC_{sw} and, consequently, that risks are estimated to be low under these circumstances

Scenario	**Step 3 $PEC_{sw;max}$ (µg/L)**	**Tier 1 RAC_{sw} (µg/L)**
D1 Ditch	0.189	0.028
D1 Stream	0.140	0.028
D3 Ditch	0.127	0.028
D4 Pond	**0.007**	0.028
D4 Stream	0.098	0.028
D5 Pond	**0.006**	0.028
D5 Stream	0.102	0.028
R4 Stream	0.569	0.028

H.1.2.2 Tier 2 effect assessment on the basis of standard and additional toxicity data

The standard aquatic macrophytes tested (Table H.2) were more sensitive than standard test algae (factor > 10) and standard test animals (factor > 1 000). Consequently, in the first instance the higher tier effect assessment focused on aquatic macrophytes.

In Table H.4 summarises the macrophyte toxicity data (good laboratory practice (GLP) and non-GLP studies in dossier and open literature) that may be used in the higher tier effect assessment. For six macrophytes valid E_rC_{50} data (based on growth rate for ecologically relevant endpoints) were available, or could be re-calculated on provided yield data. For six other aquatic macrophytes EC_{50}s on

the basis of yield endpoints could be found in the open literature, but the data provided did not allow recalculation of E_rC_{50} values.

Table H.4: Macrophytes EC_{50} values for herbicide M_A

	EC_{50} values (µg a.s./L)	**Endpoint**	**Taxonomy**
Dossier data			
Lemna minor	0.718 (7 d)	Growth rate (frond numbers)	Monocot
Lemna gibba	0.850 (7 d)	Growth rate (frond numbers)	Monocot
Myriophyllum spicatum	0.280 (14 d)	Growth rate (dry weight newly formed tissue)	Dicot
Ceratophyllum demersum	0.320 (21 d)	Growth rate (dry weight newly formed tissue)	Dicot
Open literature data			
Elodea canadensis	0.116 (21 d)	Growth rate (recalculated from yield data on new shoot length)	Monocot
Batrachium trichophyllum	0.180 (28 d)	Growth rate (recalculated from yield data on new shoot length)	Dicot
Myriophyllum aquaticum	0.220 (14 d)	Length new shoots (published data do not allow gowth rate to be calculated)	Dicot
Spirodela polyrhiza	0.230 (14 d)	Leaf area (published data do not allow gowth rate to be calculated)	Monocot
Potamogeton crispus	0.350 (21 d)	Leaf area (published data do not allow gowth rate to be calculated)	Monocot
Elodea canadensis	0.570 (21 d)	Length new shoots (published data do not allow gowth rate to be calculated)	Monocot
Ceratophyllum submersum	0.620 (21 d)	Leaf area (published data do not allow gowth rate to be calculated)	Dicot
Lemna trisulca	1.800 (7 d)	Leaf area (published data do not allow gowth rate to be calculated)	Monocot

H.1.2.3 Tier 2A: Geomean approach

For six species E_rC_{50} values (for the ecologically most relevant endpoints) are available. In the Aquatic Guidance Document is is recommended that these E_rC_{50} values are used in the effect assessment, if reported or on the basis of recalculated yield data. The geometric mean E_rC_{50} value for the top six species in Table H.4 is 0.323 µg/L. Applying an AF of 10 results in a tier 2A RAC_{sw} of 0.0323 µg/L.

H.1.2.4 Tier 2B: Species sensitivity distribution (SSD) approach

According to the AMRAP document (Maltby et al., 2010) and the Aquatic Guidance Document, when applying the SSD approach for macrophyte effect assessment, it is recommended that growth rate endpoints are used, and when growth rate estimates are not available EC_{50} values based on other sensitive endpoints should be used. For six macrophyte species E_rC_{50} values are available for the ecologically most relevant endpoints. Since according to the Aquatic Guidance Document a minimum of eight valid EC_{50} values are required to apply the SSD approach, all toxicity data mentioned in Table H.4 are used (12 species, 6 E_rC_{50} values and 6 E_yC_{50} values).

The **HC_5 value** (and 95 % confidence interval) on the basis of acute toxicity data for macrophytes (n = 12) is: **0.1065 (0.0477 – 0.1727) µg/L**. Consequently for aquatic macrophytes the **median HC_5 is 0.1065 µg/L** and the lower limit HC_5 is 0.0477 µg/L. The corresponding SSD curve is presented below (Figure H.2). The Anderson–Darling test for normality is accepted at all levels. For primary producers

the Aquatic Guidance Document recommends applying an AF of 3 to the median HC_5 of 0.1065 µg/L resulting in a **tier 2B SSD-RAC_{sw} of 0.0355 µg/L** for aquatic macrophytes. Note that this value is probably conservative since the SSD was constructed using several 'non-growth rate' endpoints.

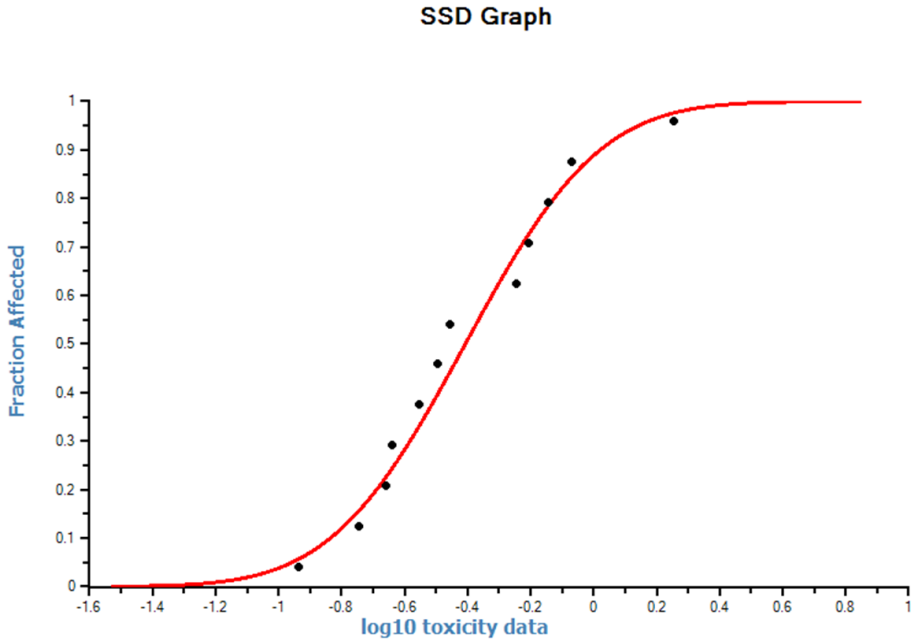


Figure H.2: Species sensitivity distribution (SSD) curve for herbicide M_A constructed with EC_{50} data for aquatic macrophytes mentioned in Table H.4

H.1.2.5 Tier 3: Micro-/mesocosm approach

An appropriate micro-/mesocosm test with a high enough diversity of aquatic macrophytes (the most sensitive taxonomic group) could not be found.

H.1.2.6 Summary effect assessment herbicide M_A

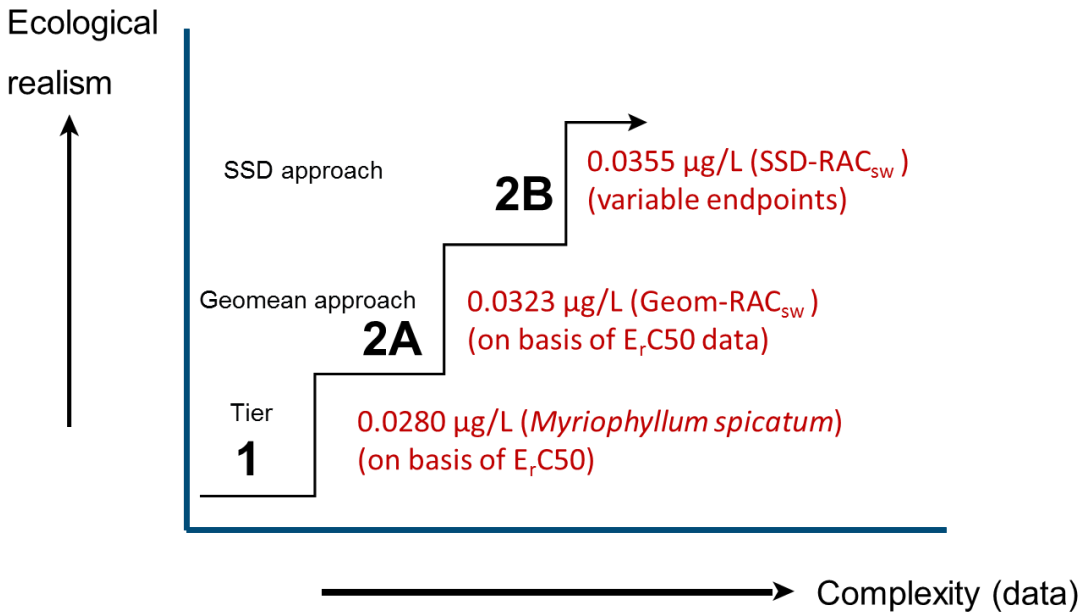


Figure H.3: Schematic presentation of the RAC_{sw} values derived on the basis of different tiers for the herbicide M_A and toxicity data for macrophytes

On the basis of the data presented above, the **tier 2B RAC_{sw}** value on the basis of macrophyte toxicity data is **0.0355 µg/L**.

H.1.2.7 Linking exposure to effects in the risk assessment

For the final risk assessment a RAC_{sw} value of 0.0355 µg/L will be used (SSD-RAC_{sw}). When revisiting the tier 1 toxicity data for all standard test species (see Table H.2) it appears that this higher tier RAC is probably protective for aquatic algae and animals. To address the aquatic risks in the first instance the SSD-RAC_{sw} value of 0.0355 µg/L has to be compared with the $PEC_{sw;max}$ (see Table H.5).

Table H.5: $PEC_{sw;max}$ values for herbicide M_A in different FOCUS scenarios and different risk-reducing measures (90 % drift-reducing nozzles). The red values indicate that the $PEC_{sw;max}$ is higher than the RAC_{sw} (0.0355 µg/L) and that under these circumstances the estimated risks are unacceptable. The black and bold values indicate that the $PEC_{sw;max}$ is lower than the RAC_{sw} and, consequently, that risks are estimated to be low under these circumstances

	$PEC_{sw;max}$ of herbicide M_A (µg/L)	
Scenario	**Step 3**	**Step 4: 90 % drift reduction**
D1 Ditch	0.189	0.115
D1 Stream	0.140	0.092
D3 Ditch	0.127	**0.013**
D4 Pond	**0.007**	**0.007**
D4 Stream	0.098	**0.010**
D5 Pond	**0.006**	**0.002**
D5 Stream	0.102	**0.011**
R4 Stream	0.569	0.569

From the results presented in Table H.5 it appears that for all the Ditch and Stream scenarios risk mitigation measures (step 4 exposure assessment) are required. Possible mitigation options for step 4 simulations are the use of drift-reducing nozzles or buffer strips.

Several scientific papers have demonstrated that in laboratory toxicity tests the effects of time-variable exposure concentrations of amino acid biosynthesis-inhibiting herbicides on the growth of *Myriophyllum spicatum* and *Lemna gibba* can best be predicted by area under the curve exposure concentrations. This indicates that for herbicide M_A and aquatic macrophytes the time-weighted average (TWA) approach is appropriate. Under these conditions the Aquatic Guidance Document offers the possibility to use the $PEC_{sw;7d\text{-}twa}$ in the risk assessment. This can be done as the EC_{50} values presented in Table H.4 are expressed in terms of mean exposure concentrations during the toxicity test. A risk assessment using $PEC_{sw;7d\text{-}twa}$ estimated for the different edge-of-field scenarios is presented in Table H.6.

Table H.6: $ECs_{w;7d\text{-}twa}$ values for herbicide M_A in different FOCUS scenarios and different risk reducing measures (90 % drift-reducing nozzles). The red values indicate that the $PEC_{sw;max}$ is higher than the SSD-RAC_{sw} (0.0355 µg/L) and that under these circumstances the estimated risks are not low. The black and bold values indicate that the $PEC_{sw;7d\text{-}twa}$ is lower than the RAC_{sw} and, consequently, that risks are estimated to be low under these circumstances

	$PEC_{sw;7d\text{-}twa}$ of herbicide M_A (µg/L)	
Scenario	**Step 3**	**Step 4: 90 % drift reduction**
D1 Ditch	0.115	0.111
D1 Stream	0.069	0.069
D3 Ditch	**0.017**	**0.002**
D4 Pond	**0.007**	**0.007**
D4 Stream	**0.005**	**0.005**
D5 Pond	**0.006**	**0.002**
D5 Stream	**0.001**	**0.001**
R4 Stream	0.079	0.079

H.1.3 Met-M_A (relevant metabolite of herbicide MA): general information

The metabolite of herbicide M_A is formed in soil. FOCUS scenario step 3 calculations result in $PEC_{sw;max}$ values in the range of 0.060 (D5 Stream) to 0.129 µg/L (D5 Pond). Corresponding $PEC_{sw;7d\text{-}twa}$ values range from 0.008 µg/L (R4 Stream) to 0.128 µg/L (D4 and D5 Pond). The following section discusses the metabolite met-M_A effect/risk assessment (> 10 % AR in soil degradation studies under laboratory conditions).

H.1.3.1 Effect/risk assessment assuming similar toxicity as parent compound (toxophore not present)

Met-M_A was formed in the soil and the toxophore was lost. No measurements of metabolite concentration were performed in the single species toxicity tests studies with the active ingredient (M_A).

As a pragmatic and conservative approach in the effect assessment for metabolites with no toxophore (see Aquatic Guidance Document, section 10.2), in first instance it can be assumed that the acute and chronic toxicity of the metabolite is equal to that of the active substance (a.s.) for all tier 1 aquatic test species (see Table H.7).

Assuming similar toxicity of the a.s. and the metabolite and by comparing the met-M_A RAC values estimated for the different standard test species (Table H.7) with the predicted step 2 $PEC_{sw;max}$ and $PEC_{sw;7d\text{-}twa}$ values for the different edge-of-field FOCUS scenarios (

Table **H.8**), it appears that the environmental risk of exposure to met-M_A cannot be excluded. However, only aquatic macrophytes are potentially at risk (step 2 $PEC_{sw;ac} > RAC_{sw}$).

Table H.7: Effect assessment for the metabolite met-M_A assuming same toxicity as the a.s. The RAC estimation on the basis of the predicted toxicity for *Myriophyllum spicatum* results in the lowest RAC estimate for Met-M_A

	Endpoint	L(E)C_x/NOEC M_A (µg/L)	Met-M_A (µg/L) assuming equal toxicity as of a.s. on a molar basis	AF	RAC Met-M_A (µg/L)
Acute risk assessment					
Oncorhynchus mykiss	96 h LC_{50}	100 000	93 333	100	933
Daphnia magna	48 h EC_{50}	> 150 000	140 000	100	1 400
Chronic risk assessment					
Pseudokirchneriella subcapitata	72 h E_rC_{50}	119	111	10	11.1
Navicula peliculosa	96 h E_rC_{50}	95	89	10	8.9
Lemna minor	7 d E_rC_{50}	0.718	0.670	10	0.067
Myriophyllum spicatum	14 d E_rC_{50}	0.28	0.26	10	**0.026**
Oncorhynchus mykiss	21 d NOEC	60 000	56 000	10	5 600
Daphnia magna	21 d NOEC	150 000	140 000	10	14 000

Table H.8: Concentration in water (µg/L) of metabolite met-M_A in different edge-of-field FOCUS scenarios (step 2 calculations). The red PEC_{sw} values are higher than the predicted RAC_{sw} of 0.026 µg/L estimated for met-M_A, so potential risks are not demonstrated to be low for any of the scenarios

	Met-M_A metabolite		
Scenario	**Step 2 $PEC_{sw;max}$ (µg/L)**	**Step 2 $PEC_{sw;7d\text{-}twa}$ (µg/L)**	**Estimates RAC_{sw} (µg/L)**
Northern Europe	0.16	0.15	0.026
Southern Europe	0.32	0.31	0.026

H.1.3.2 Effect/risk assessment on the basis of laboratory toxicity tests with *Myriophyllum spicatum* and metabolite met-MA

As the data presented in Table H.7 and
Table H.**8** reveal that macrophytes are potentially at risk, a standard laboratory toxicity test was performed with metabolite met-M_A and the potentially most sensitive standard test macrophyte *Myriophyllum spicatum*. This test resulted in a 14-d E_rC_{50} of 1600 µg/L (growth rate endpoint dry weight). Applying an AF of 10 gives an adjusted RAC_{sw} for met-M_A of 160 µg/L. This RAC is substantially higher than the step 2 $PEC_{sw;max}$ values calculated for all edge-of-field scenarios (
Table H.**8**). Consequently, aquatic macrophytes are predicted to suffer low risks when exposed to the step 2 $PEC_{sw;max}$ values for met-M_A. From the data presented in Table H.7, it already appears that for other aquatic organisms (algae, invertebrates, fish) low risks are identified.

H.2 Fungicide F_A case study

H.2.1 F_A (anilinopyrimidine): general information

Fungicide F_A, a member of the anilinopyrimidine group, is a systemic foliar broad-spectrum fungicide. It acts as an inhibitor of methionine biosynthesis and interferes in the fungal life cycle by inhibition of penetration and by disruption of mycelial growth in the plant. It has registered uses in many countries on many crops (e.g. agriculture, horticulture, viticulture).

In this case the effect/risk assessment is performed for the use of fungicide F_A in winter wheat. The use pattern shown in Table H.1 was used.

Some typical predicted exposure profiles for edge-of-field surface waters on the basis of FOCUS scenarios and models are presented in Figure H.4.

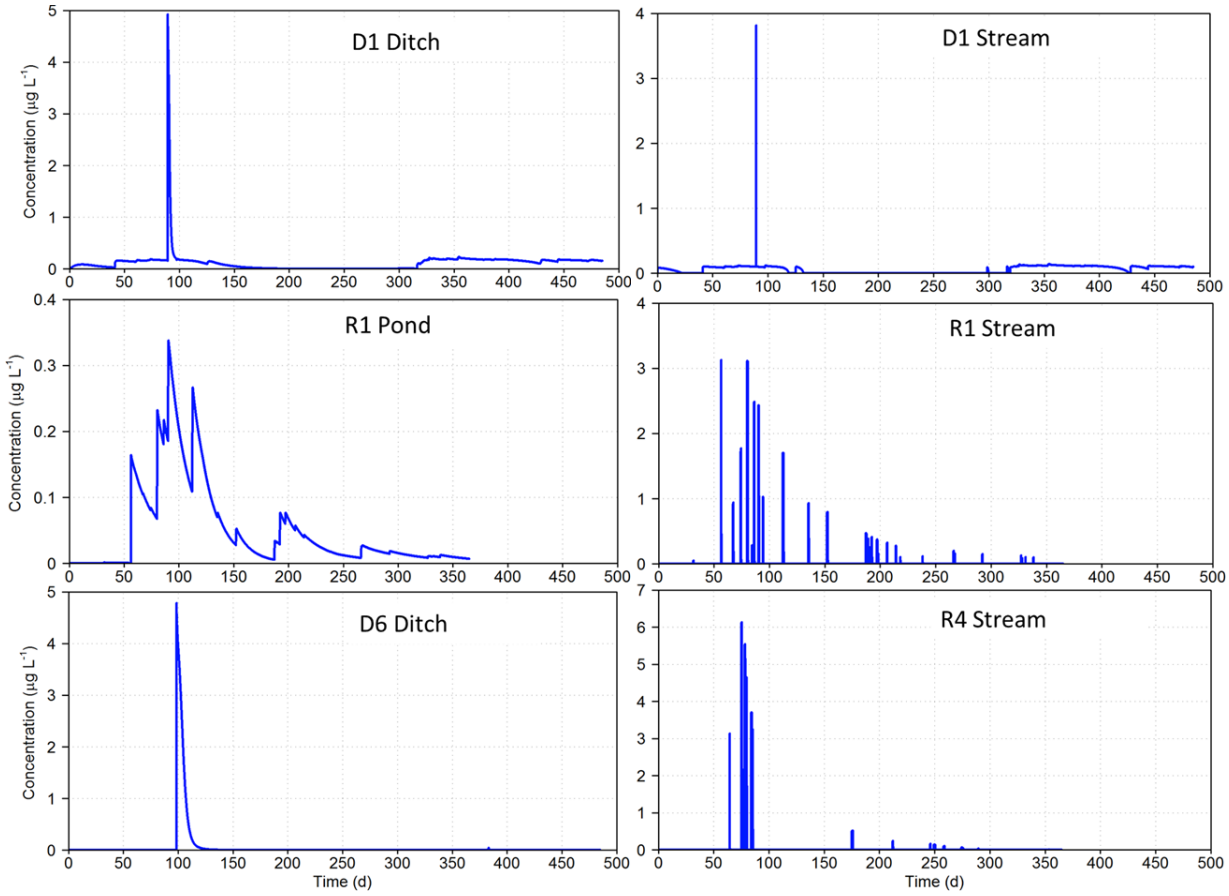


Figure H.4: Typical exposure profiles predicted for fungicide F_A in edge-of-field surface waters on the basis of FOCUS scenarios and models (step 3 calculations)

FOCUS scenario step 3 calculations result in $PEC_{sw;max}$ values in the range 0.163–5.050 µg/L. The lowest $PEC_{sw;max}$ value is calculated for D4 and D5 Pond and the highest $PEC_{sw;max}$ value for R4 Stream.

H.2.1 Tier 1 acute effect assessment

Acute toxicity data for aquatic standard test species are presented in Table H.9.

Table H.9: Acute toxicity data for aquatic standard test species and fungicide F_A

	48–96 h L(E)$_{50}$ (µg a.s./L)	**Taxonomy**	**Remark**
Americamysis bahia	18.2 (96 h)	Crustacea; Mysidae	Data requirement USA
Daphnia magna	33 (48 h)	Crustacea; Daphniidae	
Chironomus riparius	1 900 (48 h)	Insecta; Chironomidae	Preferred second arthropod
Oncorhynchus mykiss	2 500 (96 h)	Pisces; Salmonidae	
Scenedesmus subspicatus	4 600 (72 h)	Green alga	Chronic test (growth rate)

A test on a second invertebrate species is required as *Daphnia* is an order of magnitude more sensitive than algae and fish. The notifier supplied acute toxicity data for *Americamysis bahia* (data requirement in USA) and *Chironomus riparius* (preferred second test species in Aquatic Guidance Document for insecticidal mode of action). Applying an AF of 100 to the lowest toxicity value of 18.2 µg/L results in a **tier 1 acute RAC$_{sw;ac}$ of 0.182 µg/L (*Americamysis bahia*).**

H.2.2 Tier 2 acute effect assessment on the basis of standard and additional toxicity data

H.2.2.1 Tier 2A: Geomean–assessment factor approach

Acute toxicity data are available for eight crustaceans, three insects, two gastropods, four algae and one macrophyte (see Table H.10). Overall, the crustaceans are more sensitive than the other taxonomic groups. The geometric mean EC_{50} value for crustaceans is 245.9 µg/L. Applying an AF of 100 to this Geomean value (crustaceans) results in a **tier 2A Geom-RAC$_{sw;ac}$ of 2.459 µg/L**.

Table H.10: Acute toxicity data for standard and additional aquatic test species of invertebrates and fungicide F_A

	48–96 h L(E)$_{50}$ (µg a.s./L)	**Taxonomy; family**
Americamysis bahia	18.2	Crustacea; Mysidae
Daphnia magna	33	Crustacea; Daphniidae
Daphnia longispina	120	Crustacea; Daphniidae
Ceriodaphnia dubia	150	Crustacea; Daphniidae
Asellus aquaticus	210	Crustacea; Asellidae
Mesocyclops leuckarti	220	Crustacea; Cyclopidae
Halella azteca	980	Crustacea; Hyalellidae
Thamnocephalus platyurus	1 100	Crustacea; Thamnocephalidae
Chironomus riparius	1 900	Insecta; Chironomidae
Chaoborus sp.	3 500	Insecta; Chaoboridae
Cloeon dipterum	3 850	Insecta; Baetidae
Crassostrea virginia	758	Bivalvia; Ostreidae
Lymnaea stagnalis	2 900	Gastropoda; Lymnaeidae
Scenedesmus subspicatus	4 600	Green alga (endpoint growth rate)
Pseudokirchneriella subcapitata	1 200	Green alga (endpoint biomass)
Navicula pelliculosa	3 760	Diatom (endpoint cell density)
Anabaena flos-aquae	2 200	Blue-green alga (cell density)
Lemna gibba	3 200	Vascular plant (growth rate)

H.2.2.2 Tier 2B: Species sensitivity distribution (SSD) approach

As more than eight toxicity data for different taxa are available it should be explored whether the tier 2 assessment can be based on the SSD approach. The Aquatic Guidance Document states that, initially,

for a fungicide, all available aquatic toxicity data can be used to generate a SSD. In the first instance this approach is followed by constructing an SSD with all toxicity data mentioned in Table H.10. Note that the acute toxicity value for fish is not included because a higher protection level is required for aquatic vertebrates than for non-vertebrates.

The HC_5 value (and 95 % confidence interval) on the basis of acute toxicity data for all non-vertebrates ($n = 18$; see Table H.10) is 41.91 (11.03–102.44) µg/L. Consequently, for all non-vertebrate taxa the median HC_5 is 41.91 µg/L and the lower limit HC_5 is 11.03 µg/L. The corresponding SSD curve is presented below (Figure H.5). The overall fit of the curve to the toxicity data is not very good. Furthermore, the Anderson–Darling test for normality is not accepted at the 0.1 and 0.05 level. In that case the Aquatic Guidance Document recommends constructing the SSD with fewer taxonomic groups that should comprise the lowest toxicity values, but with toxicity data for as many taxonomic groups as possible, until a good fit of the SSD curve is obtained (Anderson–Darling test).

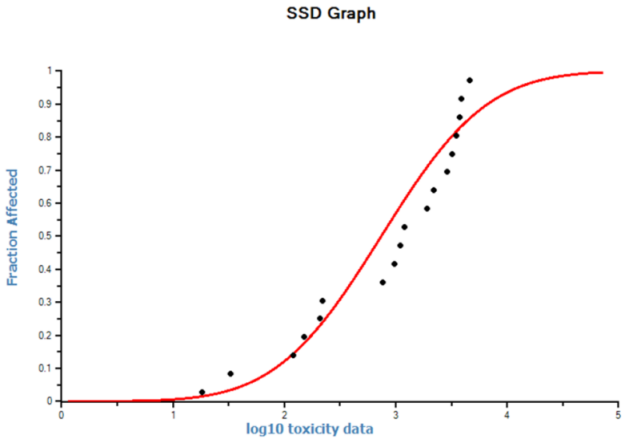


Figure H.5: Species sensitivity distribution (SSD) curve for fungicide F_A constructed with acute toxicity data of all taxa (non-vertebrates) recorded in Table H.10

Therefore, the SSD approach was also applied using the acute toxicity data for all invertebrates mentioned in Table H.10. The HC_5 value (and 95 % confidence interval) on the basis of acute toxicity data for invertebrates ($n = 13$; see Table H.10) is: 23.04 (4.08 – 66.91) µg/L. Consequently for all invertebrate taxa the median HC_5 is 23.04 µg/L and the lower limit HC_5 is 4.08 µg/L. The corresponding SSD curve is presented below (Figure H.6). The Anderson–Darling test for normality is accepted at all levels.

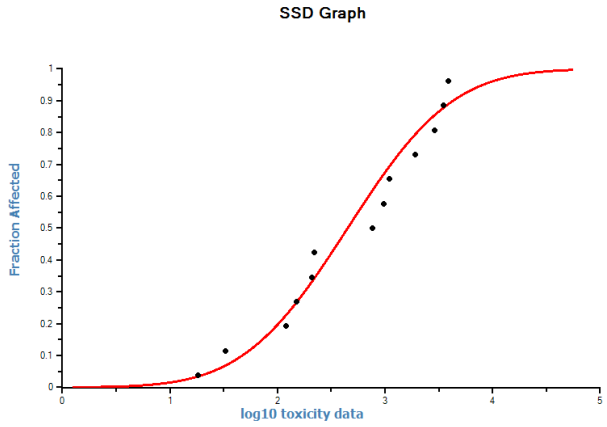


Figure H.6: Species sensitivity distribution (SSD) curve for fungicide F_A constructed with acute toxicity data of all invertebrates recorded in Table H.10

As for eight different crustaceans acute toxicity values were available, and crustaceans comprised the most sensitive invertebrates, the SSD approach was also applied to these taxa (Figure H.7). Note that these crustaceans comprise both freshwater and saltwater species. The HC_5 value (and 95 % confidence interval) on the basis of acute toxicity data for Crustacea (n = 8; see Table H.10) is 13.90 (1.68 – 41.50) µg/L. Consequently, for all Crustacea taxa the median HC_5 is 13.90 µg/L and the lower limit HC_5 is 1.68 µg/L. The corresponding SSD curve is presented below. The Anderson–Darling test for normality is accepted at all levels. The SSD curve for crustaceans and the corresponding HC_5 value is presented for illustrative purposes only. According to the recommendation in the Aquatic Guidance Document the effect assessment will be based on the SSD constructed with all invertebrates, since that curve made use of more toxicity data (n = 13) and fitted the toxicity data very well.

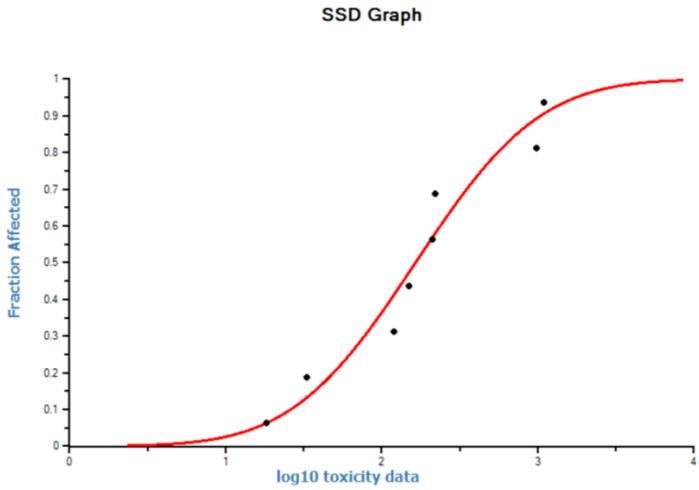


Figure H.7: Species sensitivity distribution (SSD) curve for fungicide F_A constructed with acute toxicity data of aquatic crustaceans recorded in Table H.10

The Aquatic Guidance Document recommends applying an AF of 3–6 to the median HC_5. The SSD output for invertebrates (median HC_5 of 23.04 µg/L) was selected for the tier 2B effect assessment.

Considering the criteria mentioned in section 8.4.4 of the Aquatic Guidance Document we decided to select an AF of 6 to be applied to the median HC_5 (resulting in a value of 3.84 µg/L) since:

1. Application of the proposed AF of 6 to the median HC_5 will not result in a lower SSD-$RAC_{sw;ac}$ than the tier 1 $RAC_{sw;ac}$ of 0.182 µg/L and the tier 2A $RAC_{sw;ac}$ of 2.46 µg/L.

2. Although the SSD curve constructed with all invertebrates (see Figure H.6) fits the toxicity data in the tail of the SSD curve fairly well, overall the crustaceans tested were more sensitive than the other invertebrates tested, hence indicating selecting a higher AF.

3. The lower limit HC_5 (4.08 µg/L) derived from the invertebrate SSD curve was less than one-third of the median HC_5 (23.04 µg/L), hence giving indications to select a higher AF.

4. The SSD curve is not very steep (more than a factor of 100 between lowest and highest $L(E)C_{50}$), hence indicating that a AF in the lower range may be selected.

5. The acute to chronic ratio for at least one standard test species (*Chironomus*) and the additional test species *Asellus aquaticus* is > 10 (see Table H.9 and Table H.12), hence warranting an AF in the higher range.

Applying an AF of 6 to the median HC_5 value of 23.04 µg/L given above for invertebrates results in a **tier 2B SSD-$RAC_{sw;ac}$ of 3.84 µg/L** (on the basis of aquatic invertebrates).

H.2.2.3 Tier 3: Micro-/mesocosm approach

One valid mesocosm study is available that was conducted under GLP. It was classified as a Ri-2 study (reliability index, see section 9.3.3).[2] To the mesocosms a formulation with fungicide F_A was applied three times at weekly intervals (in June). The last sampling of organisms took place 10 weeks after the last application. At the second and third application not all a.s. had disappeared from the water column so that the peak concentration of fungicide F_A was measured immediately after the third application. The overall dissipation DT_{50} of the a.s. in water was estimated to be 9–11 days. Most macroinvertebrate taxa that were affected by F_A application showed the maximum effect in the weeks after the last application. The overall summary of effect classes observed for several categories of endpoints in the GLP mesocosm study are presented below (Table H.11). For at least 10 invertebrate populations (including 8 crustaceans) a consistent concentration–response relationship for negative effects and corresponding NOEC/LOEC values could be derived. For three other invertebrate taxa treatment-related increases were observed. For the purpose of this example it is assumed that the MDDs are low enough. Only a few populations of phytoplankton taxa and insect taxa showed a minor treatment-related response.

Table H.11: Overall summary of effect classes observed for several categories of endpoints in the outdoor mesocosm study receiving three applications of fungicide F_A

	Treatment levels (µg a.s./L)				
Nominal concentration	**2.0**	**6.0**	**18.0**	**54.0**	**162.0**
Measured peak concentration	**4.0**	**9.1**	**23.6**	**68.5**	**210.0**
Highest 7-d TWA concentration	**3.3**	**7.3**	**19.4**	**58.2**	**170.4**
Highest 21-d TWA concentration	**2.5**	**5.4**	**14.8**	**43.6**	**125.8**
Population responses					
Macroinvertebrates					
Gammarus pulex (Crustacea)	1	1	2	3B	5B

2 No MDD calculations were preformed in this study; they would, however, be needed before the study could be used for a regulatory decision. However, the study is now included in order to illustrate how to link exposure and effect assessment.

	Treatment levels (µg a.s./L)				
Asellus aquaticus (Crustacea)	1	2	3A	5B	5B
Stylaria lacustris (Oligochaeta)	1	1	1	2↑	3A↑
Dugesia lugubris (Tricladida)	1	1	1	1	2
Bythinia tentaculata (Gastropoda)	1	1	1	1	3A
Lymnaea stagnalis (Gastrpoda)	1	1	1	2↑	3A↑
Chironomidae (Insecta)	1	1	1	1	3A↑
Ischnura elegans	1	1	1	1	2
Chaoborus obscuripes	1	1	1	1	2
Ephemeroptera (Insecta)	1	1	1	1	2
Zooplankton					
Daphnia longispina (Crustacea)	1	2	3A	3A	3A
Simocephalus vetulus (Crustacea)	1	1	2	3A	3A
Alona sp.	1	1	2	3A	3A
Nauplia sp.	1	1	1	3A	5A
Cyclopoida	1	2	2	3B	5B
Calanoida	1	1	2	3A	5A
Copepoda	1	2	2	3B	5B
Lecane arcuata (Rotifera)	1	1	2	3A	3A
Keratella quadrata (Rotifera)	1	1	1	2↑	3A↑
Phytoplankton					
Volvox (green alga)	1	1	1	2↑	2↑
Scenedesmus aculeolatus (green alga)	1	1	1	2	3A
Phytoplankton chlorophyll-a	1	1	1	1	2
Community responses					
Macroinvertebrates	1	2	3A	5A	5B
Zooplankton	1	1	3A	3A	5A
Phytoplankton	1	1	1	2	2
Physicochemical measurements	1	1	1	1	1
Most sensitive endpoint	**1**	**2**	**3A**	**5B**	**5B**

↑Treatment-related increase in abundance; a.s, active substance.

To address the ecological threshold option the effect class 1 concentrations of 4.0 µg/L (measured peak concentration) may be used in the effect assessment by applying an AF of 2. This procedure results in a **tier 3 acute ETO-RAC$_{sw;ac}$ of 2.00 µg/L**.

Alternatively, to address the ecological threshold option the effect class 2 concentrations of 9.1 µg/L (measured peak concentration) may be used in the effect assessment by applying an AF of 2 to 3. The AF of 3 is selected as the study was classified as an Ri-2 study. This procedure results in a **tier 3 acute ETO-RAC$_{sw;ac}$ of 3.03 µg/L**. In the effect assessment this last value is selected as ETO-RAC$_{sw;ac}$, since the decline of the most sensitive species *Asellus* was less than 25 % for the effect class 2 concentration. In addition, an effect class 2 concentration is a better indicator for the effect threshold than an effect class 1 concentration, since the actual threshold will always be higher than the NOEC of most sensitive endpoint.

The lentic mesocosm study contained two populations of macro-crustaceans, namely *Asellus aquaticus* and *Gammarus pulex* (the most sensitive taxonomic group on the basis of laboratory toxicity data) that

can be classified as relatively vulnerable, in the sense that recovery potential is low in isolated lentic mesocosms. Furthermore, laboratory toxicity data and responses in the mesocosm study revealed that representatives of insects, worms and snails (including univoltine taxa) were relatively insensitive. To address the ecological recovery option the effect class 3A concentration of 23.6 µg/L is used in the effect assessment by applying an AF of 3–4. The AF of 4 is selected as the study was classified as an Ri-2 study. This procedure results in a **tier 3 acute ERO-RAC$_{sw;ac}$ of 5.90 µg/L**. Since the duration of the pulse exposure tested in the mesocosms was realistic to worst case relative to the duration of the highest predicted pulse in the FOCUS scenarios (except perhaps R1 Pond; see Figure H.4), this acute ERO-RAC$_{sw;ac}$ value of 5.90 µg/L can be used in the acute risk assessment for most FOCUS scenarios. Note that the PEC$_{sw;max}$ in the R1 Pond scenario (0.338 µg/L) is lower than the tier 1 RAC$_{sw;ch}$ (0.75 µg/L; see below in chronic risk effect assessment), so that potential risks to aquatic organisms of exposure to fungicide F_A are likely negligible.

Recommendation regarding the use of ERO-RAC (based on the effect class 3A concentration for *Asellus*). In the climatic zone of concern this species is bivoltine and reproduces in spring and late summer. This effect assessment considers the fungicide use in spring and early summer so that recovery of *Asellus* effectively can take place (allowing the use of the ERO-RAC derived). However, if in other crops the use of this fungicide is in late summer, we would suggest not using the ERO RAC, since in autumn recovery of *Asellus* within eight weeks cannot be guaranteed.

Since in the mesocosm study fish was not present it should be checked whether the ETO-RAC$_{sw;ac}$ and the ERO-RAC$_{sw;ac}$ are sufficiently protective for aquatic vertebrates. In the tier 1 dataset (Table H.9) the reported 96-hour LC$_{50}$) for *Oncorhynchus mykiss* is 2 500 µg/L. Applying an AF of 100 results in a tier 1 RAC$_{sw;ac}$ for aquatic vertebrates of 25 µg/L. Since this value is higher than the ETO-RAC$_{sw;ac}$ and the ERO-RAC$_{sw;ac}$ mentioned above, the acute risk assessment can be based on the results of the mesocosm study.

H.2.3 Fungicide F_A chronic effect assessment

H.2.3.1 Tier 1 chronic effect assessment

Chronic toxicity data for standard test species are presented in Table H.12.

Table H.12: Chronic toxicity data for aquatic standard and additional test species and fungicide F_A

	L(E)$_x$/NOEC (µg a.s./L)	**Taxonomy**	**Remark**
Standard test species			
Daphnia magna	7.5 (21 d NOEC; reproduction)	Crustacea	
Chironomus riparius	140 (28 d EC10)	Insecta	Water–sediment study
Pimephales promelas	485 (36 d NOEC)	Fish	
Scenedesmus subspicatus	1200 (72 h E_rC_{50})	Green alga	
Additional test species			
Asellus aquaticus	12.1 (28 d NOEC; mortality)		

Applying an AF of 10 to the lowest toxicity value of 7.5 µg/L results in a **tier 1 chronic RAC of 0.75 µg/L** (*Daphnia magna*).

H.2.3.2 Tier 2 chronic effect assessment on the basis of standard and additional toxicity data

An additional chronic toxicity value is available for the invertebrate *Asellus aquaticus*. Since the NOEC values for the crustaceans *Daphnia* and *Asellus* concern different endpoints, the Geomean

approach cannot be applied in the chronic risk assessment. For the SSD approach not enough toxicity data are available.

H.2.3.3 Tier 3: The model ecosystem approach

Under the condition that the exposure regime of the mesocosm study is relatively the worst case for the long-term exposure regime in the field this study might also be used for the long-term risk assessment by comparing the tier 3 RAC (based on effect class 2 expressed in terms of peak concentrations in the test system) with the $PEC_{sw;max}$ (for results see Table H.11).

The exposure regime simulated in the mesocosm study (three weekly applications) is considered realistic to worst case for all FOCUS edge-of-field scenarios except for R1 Pond (see Figure H.4). Consequently, effect class concentrations expressed in terms of peak concentrations can be used to derive an ETO-$RAC_{sw;ch}$ and ERO-$RAC_{sw;ch}$ to be used in the chronic risk assessment for all scenarios except R1 Pond. In fact, in these scenarios (except R1 Pond) the ETO-$RAC_{sw;ac}$ and ETO-$RAC_{sw;ch}$, as well as the ERO-$RAC_{sw;ac}$ and the ERO-$RAC_{sw;ch}$, are similar **(tier 3 acute ETO-$RAC_{sw;ch}$ is 3.03 µg/L and the tier 3 acute ERO-$RAC_{sw;ch}$ is 5.90 µg/L)**. Note, that in the R1 Pond scenario the $PEC_{sw;max}$ (0.338 **µg**/L) is lower than the tier 1 $RAC_{sw;ch}$ (0.75 µg/L) so that in this scenario chronic risks are not triggered.

Since in the mesocosm study, fish were not present it should be checked whether the ETO-$RAC_{sw;ch}$ and the ERO-$RAC_{sw;ch}$ are sufficiently protective for aquatic vertebrates. In the tier 1 chronic dataset (Table H.12) the reported 36-day NOEC for *Pimephales promelas* is 275 µg/L. Applying an AF of 10 results in a tier 1 $RAC_{sw;ch}$ for aquatic vertebrates of 25.5 µg/L. Since this value is higher than the ETO-$RAC_{sw;ac}$ and the ERO-$RAC_{sw;ac}$ mentioned above, the acute risk assessment can be based on the results of the mesocosm study.

H.2.4 Summary of acute and chronic effect assessment fungicide F_A

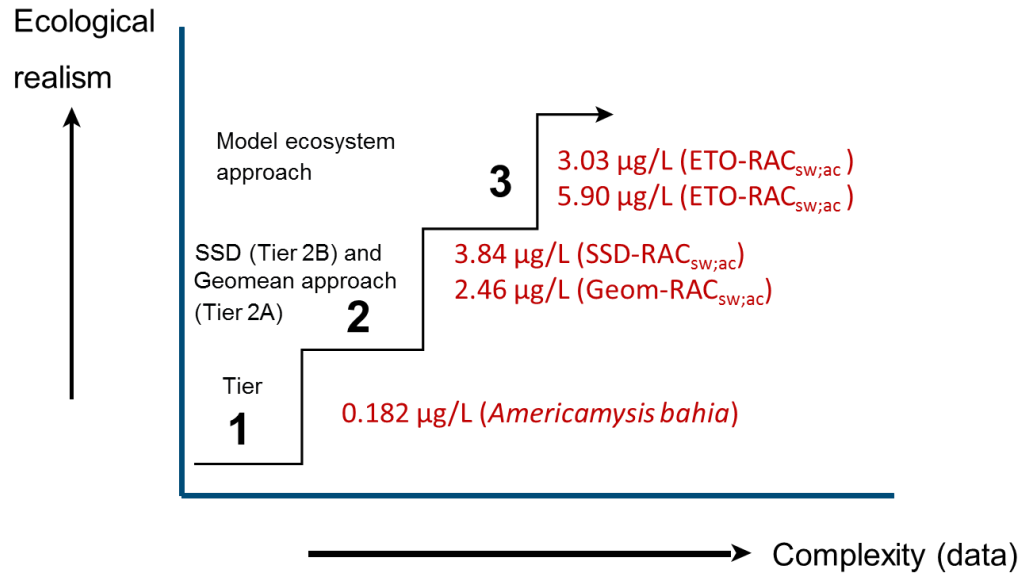


Figure H.8: Schematic presentation of the $RAC_{sw;ac}$ values derived on the basis of different tiers for the fungicide F_A in the acute effect assessment (see above the recommendations regarding the use of ERO-RAC)

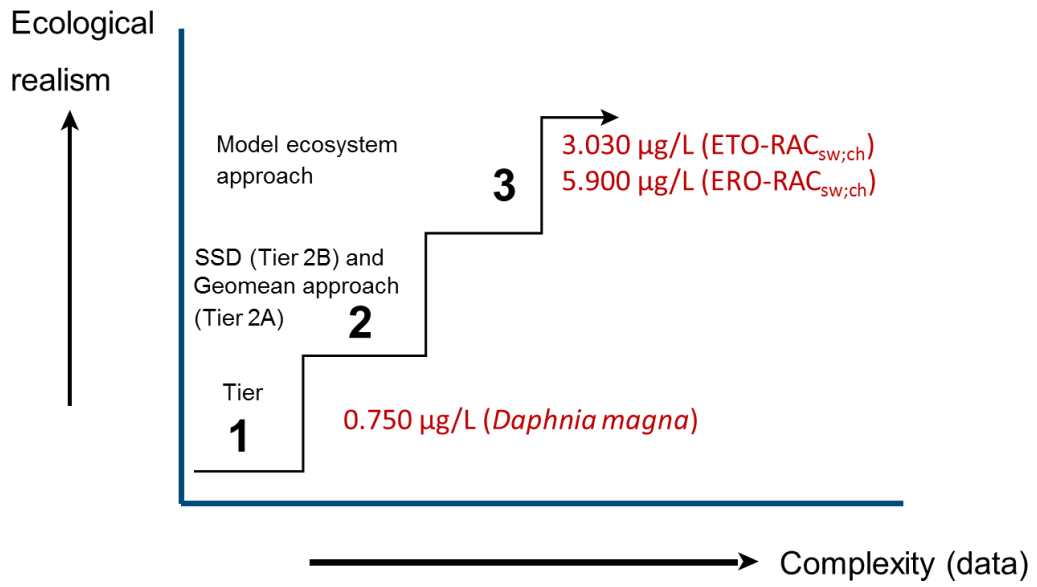


Figure H.9: Schematic presentation of the $RAC_{sw;ch}$ values derived on the basis of different tiers for the fungicide F_A in the chronic effect assessment. In this case the ETO-RAC and ERO-RAC values should always be compared with the $PEC_{sw;max}$

H.2.5 Linking exposure to effects in the risk assessment

As final $RAC_{sw;ac}$ and $RAC_{sw;ch}$ values, the tier 3 results are adopted, i.e. the ETO-RAC_{sw} of 3.030 µg/L and an ERO-RAC_{sw} of 5.900 µg/L. This can be done because, except for the R1 Pond scenario, the exposure regime tested in the mesocosms is more or less realistic to worst case relative to the predicted exposure profiles for the different FOCUS scenarios. The tier 3 RAC values have to be compared with the $PEC_{sw;max}$ values as calculated for relevant FOCUS scenarios. For the ETO-RAC_{sw} and ERO-RAC_{sw} this is done in Table H.13.

Table H.13: $PECs_{w;max}$ values for fungicide F_A in different FOCUS scenarios (step 3). The red values indicate that the $PEC_{sw;max}$ is higher than the ERO-$RAC_{sw;ac}$ and/or ETO-$RAC_{sw;ac}$ and that under these circumstances the estimated risks are unacceptable. The black and bold values indicate that the $PEC_{sw;max}$ is lower than the ERO-$RAC_{sw;ac}$ (threshold option) and/or ETO-$RAC_{sw;ac}$ (recovery option) and, consequently, that risks are estimated to be acceptable under these circumstances

	Threshold option		**Recovery option**[(a)]	
Scenario	**$PEC_{sw;max}$ (µg/L)**	**ETO-$RAC_{sw;ac}$ ETO-$RAC_{sw;ch}$ (µg/L)**	**$PEC_{sw;max}$ (µg/L)**	**ERO-$RAC_{sw;ac}$ ERO-$RAC_{sw;ch}$ (µg/L)**
D1 Ditch	5.050	3.030	**5.050**	5.090
D1 Stream	3.350	3.030	**3.350**	5.090
D2 Ditch	4.923	3.030	**4.923**	5.090
D2 Stream	4.345	3.030	**4.345**	5.090
D3 Ditch	4.740	3.030	**4.740**	5.090
D4 Pond	**0.163**	3.030	**0.163**	5.090
D4 Stream	3.692	3.030	**3.692**	5.090
D5 Pond	**0.165**	3.030	**0.165**	5.090
D5 Stream	3.256	3.030	**3.256**	5.090
D6 Ditch	4.654	3.030	**4.654**	5.090
R1 Pond	**0.206**	0.75[(a)]	**0.206**	0.75[(a)]
R1 Stream	3.124	3.030	**3.124**	5.090
R3 Stream	4.388	3.030	**4.388**	5.090

	Threshold option		Recovery option[(a)]	
R4 Stream	3.124	3.030	**3.124**	5.090

(a): Tier 1 $RAC_{sw;ch}$ since exposure regime of mesocosm test is not realistic to worst case relative to exposure profile in R1 Pond.

It appears from the data in Table H.13 that, when adopting the threshold option in the risk assessment (ETO-RAC_{sw}), only the exposure profiles predicted for FOCUS scenarios D4 Pond, D5 Pond and R1 Pond will cause negligible risks. In contrast, when adopting the recovery option (ERO-RAC_{sw}) in all scenarios unacceptable risks are not triggered.

H.3 Insecticide I_N case study

H.3.1 I_N (neonicotinoid): general information

Insecticide I_N is a neonicotinoid insecticide. The neonicotinoids are a class of insecticides with a common mode of action affecting the central nervous system of insects, causing paralysis and death. Neonicotinoids block a specific neural pathway that is more abundant in insects than in warm-blooded animals. They bind at a specific site, the postsynaptic nicotinic acetylcholine receptor. As a group they are effective against sucking insects such as aphids, but also chewing insects such as Coleoptera and some Lepidoptera. Insecticide I_N is used in a wide range of different crops, including apples, tomatoes, sugar beet and maize. Because of its systemic properties it can be applied as a seed treatment, but it is also applied via spraying and drip irrigation in greenhouses.

In this case the effect/risk assessment is performed for the use of insecticide I_N in apples according to the use pattern described in Table H.1. Some typical predicted exposure profiles for edge-of-field surface waters on the basis of FOCUS scenarios and models are presented in Figure H.10.

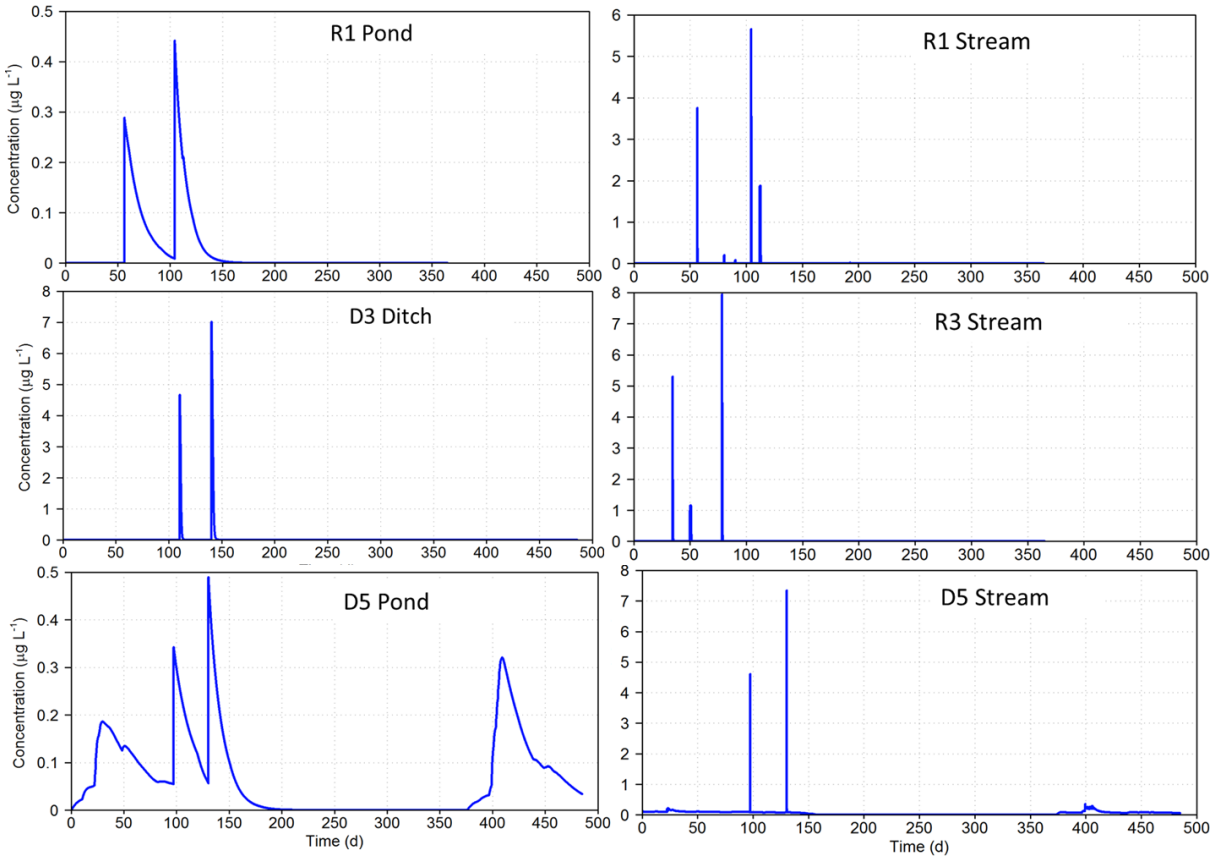


Figure H.10: Typical exposure profiles predicted for edge-of-field surface waters on the basis of FOCUS scenarios and models

FOCUS scenario step 3 calculations result in $PEC_{sw;max}$ values in the range of 0.441–7.940 µg/L. The lowest $PEC_{sw;max}$ value is calculated for R1 Pond and the highest $PEC_{sw;max}$ value for R3 Stream. Considering 90 % drift-reducing nozzles (step 4 calculations) result in $PEC_{sw;max}$ values in the range of 0.044 µg/L (R1 Pond) to 1.299 µg/L (D4 Stream).

H.3.2 I_N (neonicotinoid) acute effect assessment

H.3.2.1 Tier 1 acute effect assessment

Acute toxicity data for standard test species are presented in Table H.14.

Table H.14: Acute toxicity data for aquatic standard test species and insecticide I_N

	48–96 h $L(E)_{50}$ (µg a.s./L)	**Taxonomy**	**Remark**
Daphnia magna	37397 (48 h)	Crustacea; Daphniidae	
Chironomus tentans	5.5 (96 h)	Insecta; Chironomidae	Preferred second arthropod
Americamysis bahia	38.2 (96 h)	Crustacea; Mysidae	Data requirement USA
Oncorhynchus mykiss	211000 (96 h)	Pisces; Salmonidae	
Scenedesmus subspicatus	> 10 000 (72 h)	Green alga	Chronic test

The insect *Chironomus* and the macrocrustacean *Americamysis* are several orders of magnitude more sensitive than other standard test species (including *Daphnia magna*).

By applying an AF of 100 to the lowest toxicity value of 5.5 µg/L for *Chironomus tentans* the **tier 1 $RAC_{sw;ac}$ becomes 0.055 µg/L**.

H.3.2.2 Tier 2 acute effect assessment on the basis of standard and additional toxicity data

The tier 1 data indicate that representatives of Crustacea (*Americamysis*) and Insecta (*Chironomus*) are several orders of magnitude more sensitive than the other standard test species. The higher tier effect assessment on the basis of additional laboratory toxicity data will in the first instance focus on aquatic arthropods (crustaceans and insects).

H.3.2.2.1 Tier 2A: Geomean—assessment factor approach

Sufficient toxicity data are available to go directly to tier 2B (SSDs); however, in order to illustrate all the different approaches, the Geomean approach is shown below.

Fourteen acute toxicity data are available for crustaceans (see Table H.15). These reveal a high variability in $L(E)C_{50}$ values (ranging from 1 to 37 397 µg/L). A remarkable phenomenon is that species within the same family (Daphniidae) rank low (*Ceriodaphnia dubia*) and high (*Daphnia magna*) in the $L(E)C_{50}$ range.

Table H.15: Acute toxicity data for standard and additional aquatic test species of Crustacea and insecticide I_N

	48–96 h $L(E)_{50}$ (µg a.s./L)	**Taxonomy; family**
Daphnia magna	37 397	Crustacea; Daphniidae
Gammarus fossarum	1 335	Crustacea; Gammaridae
Chydorus ovalis	832	Crustacea; Chydoridae
Palaemonetes paludosus	417.1	Crustacea; Palaemonidae
Artemia salina	329.2	Crustacea; Artemiidae
Hyalella azteca	119	Crustacea; Hyalellidae
Asellus aquaticus	75.7	Crustacea; Asellidae
Americamysis bahia	38.2	Crustacea; Mysidae
Gammarus roeseli	18.3	Crustacea; Gammaridae
Gammarus pulex	10	Crustacea; Gammaridae
Cypridopsis vidua	5.2	Crustacea; Cyprididae
Ceriodaphnia dubia	3	Crustacea; Daphniidae

	48–96 h $L(E)_{50}$ (µg a.s./L)	Taxonomy; family
Cypretta nigra	2.07	Crustacea; Cyprididae
Ilyocryptus brevidentatus	1	Crustacea; Ilyocryptidae

The geometric mean $L(E)C_{50}$ for all crustaceans is 61.70 µg/L.

Fifteen acute toxicity data are available for insect taxa (see Table H.16). It appears that in particular EPT taxa (Ephemeroptera, Plecoptera and Trichoptera) are amongst the most sensitive species.

Table H.16: Acute toxicity data for standard and additional aquatic test species of Insecta and insecticide I_N

	48–96 h $L(E)_{50}$ (µg a.s./L)	Taxonomy; Family
Chaoborus flavicans	284	Insecta; Chaoboridae (Nematocera)
Plea minutissima	50.6	Insecta; Pleidae (Heteroptera)
Aedes communis	44	Insecta; Culicidae (Nematocera)
Sialis lutaria	35.9	Insecta; Sialidae (Megaloptera)
Notonecta glauca	18.2	Insecta; Notonectidae (Heteroptera)
Corixa striata	10.8	Insecta; Corixidae (Heteroptera)
Cloeon dipterum	8.49	Insecta; Baetidae (Ephemeroptera)
Simulium costatum	8.1	Insecta; Simuliidae (Nematocera)
Chironomus tentans	5.5	Insecta; Chironomidae (Nematocera)
Cheumatopsyge lepida	4.98	Insecta; Hydropsychidae (Trichoptera)
Rhaptobaetopus tenellus	4.2	Insecta; Baetidae (Ephemeroptera)
Limnephilus elegans	1.79	Insecta; Limnephilidae (Trichoptera)
Baetis rhodani	1.77	Insecta; Baetidae (Ephemeroptera)
Nemoura cineria	1.02	Insecta; Nemouridae (Plecoptera)
Caenis robusta	0.65	Insecta; Caenidae (Ephemeroptera)

The geometric mean $L(E)C_{50}$ for all insects is 8.41 µg/L.

As the geometric mean ($L(E)C_{50}$ for insects is lower than that for crustaceans, the value for insects is selected to derive the Geom-$RAC_{sw;ac}$ by applying an AF of 100, resulting in a **tier 2A Geom-$RAC_{sw;ac}$ of 0.084 µg/L**.

Note that more than eight toxicity data for arthropods (as well as for crustaceans and insects separately) are available. Consequently, according to the recommendations in the Aquatic Guidance Document, it will be the SSD approach rather than the Geomean approach that will be applied in the risk assessment.

H.3.2.2.2 Tier 2B: Species sensitivity distribution (SSD) approach

As the Geom-$RAC_{sw;ac}$ was based on acute toxicity of insect taxa, and the geometric mean $L(E)C_{50}$ for insects was considerably lower than that for crustaceans, the toxicity data reported in Table H.16 are used to construct an SSD curve for aquatic insects. The corresponding SSD curve is presented in Figure H.11. The Anderson–Darling test for normality is accepted at all levels. The HC_5 value (and 95 % confidence interval) on the basis of acute toxicity **data for insects** (n=15) is 0.530 (0.123–1.346) µg/L. Consequently for insect taxa the **median HC_5 is 0.530 µg/L** and the **lower limit HC_5 is 0.123 µg/L**. The Aquatic Guidance Document recommends to apply an AF of 3–6 to the median HC_5 of 0.530 µg/L (resulting in values of 0.088–0.177 µg/L).

Alternatively, the acute $L(E)C_{50}$ values for aquatic arthropods (crustaceans and insects) listed in Table H.15 and Table H.16 are used to construct SSD curve and to calculate HC_5 values. The corresponding

SSD curve for all arthropods is presented below (Figure H.12). The Anderson–Darling test for normality is accepted at all levels. The HC_5 value (and 95 % confidence interval) on the basis of acute toxicity data for **all aquatic arthropods** ($n = 29$) is: 0.326 (0.077–0.942) µg/L. Consequently for arthropod taxa the **median HC_5 is 0.326 µg/L** and the **lower limit HC_5 is 0.077 µg/L**. The Aquatic Guidance Document recommends applying an AF of 3–6 to the median HC5 of 0.326 µg/L (resulting in values of 0.054–0.109 µg/L).

Considering the criteria mentioned in section 8.4.4 of the Aquatic Guidance Document we decided to select an AF of 4.5 (mean of 3 and 6) to be applied to the median HC_5 derived for arthropods, because:

1. The lower limit HC_5 (0.077 µg/L) derived from the arthropod SSD curve was less than one-third of the median HC_5 (0.326 µg/L), hence indicating selecting a higher AF.

2. The available toxicity data comprise several different genera/families/orders of the potential sensitive taxonomic groups, including EPT (Ephemeroptera, Plecoptera, Trichoptera), taxa that have been shown to be particularly sensitive to other neonicotinoids. Thus, a lower AF in the proposed range may be selected.

3. The tier 1 $RAC_{sw;ac}$ is 0.055 µg/L and the Geomean tier 2A $RAC_{sw;ac}$ is 0.084 µg/L; applying the AF in the higher range would result in a tier 3 SSD-$RAC_{sw;ac}$ of 0.039 µg/L (arthropod SSD) or 0.062 µg/L (insect SSD), thus indicating selecting a lower AF within the range.

4. The SSD curves for insects and arthropods fit the toxicity data well, and also in the tail of the SSD curve, hence giving no indications of whether to select a higher or lower AF.

5. The SSD curves are not very steep (more than a factor of 100 between the lowest and highest $L(E)C_{50}$), hence indicating that an AF in the lower range may be selected.

6. The acute to chronic ratio for several arthropod species (*Daphnia magna*, *Hyalella azteca*, *Asellus aquaticus*, *Chaoborus flavicans*, *Plea minutissima*, *Sialis lutaria*, *Baetis rhodani*, *Caenis robusta*) is larger than 10, hence warranting an AF in the higher range (see Tables H.15, H.16, H.19 and H.20).

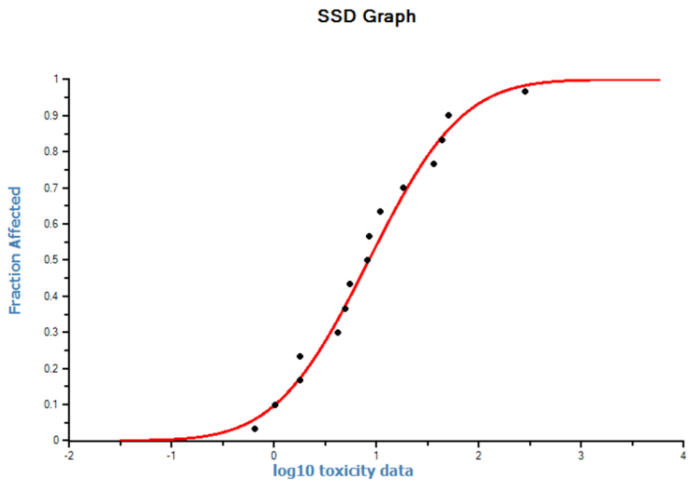


Figure H.11: Species sensitivity distribution (SSD) curve for insecticide I_N constructed with acute toxicity data from aquatic insects

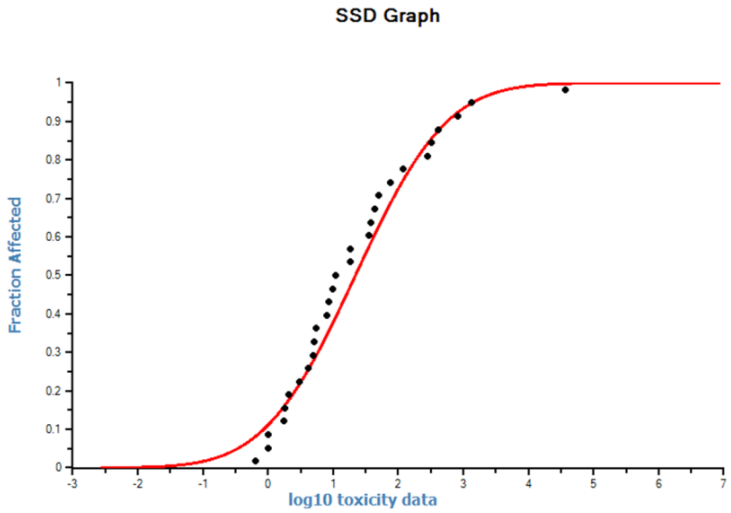


Figure H.12: Species sensitivity distribution (SSD) curve for insecticide I_N constructed with acute toxicity data from aquatic arthropods (insects and crustaceans)

Applying an AF of 4.5 to the median HC_5 (0.326 µg/L) value for all arthropods results in a **tier 2B SSD-$RAC_{sw;ac}$ of 0.072 µg/L** (on the basis of all arthropods). According to the recommendation in the Aquatic Guidance Document the effect assessment will be based on the SSD constructed with all arthropods, since that curve made use of more toxicity data (n = 29) and fitted the toxicity data very well.

H.3.2.3 Tier 3 micro-/mesocosm approach

An experimental stream mesocosm study for a period of six months post treatment is available. In this study insecticide I_N was applied once in spring. In the experimental streams, the dissipation DT_{50} was set to 15 days. Water was circulated in the streams so that the temporal concentration profile reflected the worst case situation in a pond. Sufficient numbers of populations of insects and crustaceans were present in high densities in the experimental stream study, including 22 potentially sensitive and vulnerable insect taxa as summarised in Table H.17. Note: If the ecological threshold option is considered, only the sensitivity of investigated species is relevant; in the event that the ecological recovery option is considered, the investigated species must not only be sensitive but also vulnerable (long generation time) so that unrealistically short recovery times are not obtained. Consistent concentration–response relationships and corresponding NOEC/LOEC values could be derived for 15 taxa (Table H.17).

The overall summary of effect classes observed for several endpoints in the outdoor stream mesocosm study are presented below (Table H.17). To derive RACs, the nominal concentrations are close to the measured peak concentrations, and thus are appropriate if the exposure regime in the mesocosm study is realistic to worst case when compared with the predicted exposure profile in the edge-of-field surface water of concern. If not, the two-day TWA concentrations should be used in the acute effect assessment (and as 'C' in the effect class concentrations). Also in the latter case, the $RAC_{sw;ac}$ should be compared with the FOCUS scenario step 3 $PEC_{sw;max}$.

MDDs are calculated for the NOECs of each endpoint (if it is possible to calculate them) and listed in Table H.17 according to the method presented in Appendix F.

In order to investigate the treatment-related responses at the community level three different methods were applied (see Table H.17 below community assessment analysis 1–3). The obtained NOEC varies and hence applying several different methods when evaluating the results from a mesocosm is

recommended. The PPR Panel cannot recommend a specific method as being most suitable; however, we foresee that in the future further guidance can be provided on the advantages of the various methods.

Table H.17: Overall summary of effect class responses observed for several endpoints in the outdoor experimental stream study treated with the insecticide I_N. Within each category, the most sensitive population/community level endpoint was selected. The effect class concentrations are expressed in terms of the nominal treatment concentrations (but effect classes expressed in terms of 7-day and 21-day TWA are also reported). Note that for convenience only a single MDD value is presented for each taxon. The MDD should be reported for each taxon on each sampling. The MDD NOEC presented below is the MDD value corresponding with the lowest NOEC during the period that effects are demonstrated

Concentration	**Treatment level (µg/L)**					
Nominal	0.1	0.6	3.2	20	100	600
7-d TWA	0.08	0.43	2.31	14.4	80	430
21-dTWA	0.04	0.23	1.67	7.6	40	230

Concentration / Endpoint	600	100	20	3,2	0,6	0,1	MDD NOEC	Insect	
Taxa									
Aeshna sp.	0	0	0	0	0	0	-	+	
Asellus aquaticus	3A	3A	3A	1	1	1	39		
Atherix sp.	0	0	0	0	0	0	-	+	
Baetis macani	5A	5A	5A	5A	5A	1	86	+	1)
Chironomidae	3A	3A	3A	1	1	1	32	+	
Cloeon dipterum	5A	5A	5A	3A	3A	1	48	+	
Dugesia sp.	3A	2	2	1	1	1	68		
Dytiscidae	0	0	0	0	0	0	-	+	
Ephemera vulgata	0	0	0	0	0	0	-	+	
Erpobdella	0	0	0	0	0	0	-		
Gammerus sp.	0	0	0	0	0	0	-		
Gerris gibbifer	0	0	0	0	0	0	-	+	
Helobdella sp.	0	0	0	0	0	0	-		
Hydrometra stagnorum	3A	3A	3A	1	1	1	32	+	
Hydrophilidae	0	0	0	0	0	0	-	+	
Hydropsyche angustipennis	0	0	0	0	0	0	-	+	
Ischnura elegans	3A	3A	3A	3A	3A	1	32	+	
Leucorrhinia sp.	0	0	0	0	0	0	-	+	
Libellula quadrimaculata	0	0	0	0	0	0	-	+	
Limnephilus lunatus	3A	3A	3A	3A	3A	1	69	+	
Lymnea sp.	1	1	1	1	1	1	42		
Nemoura cinerea	5A	5A	5A	5A	5A	1	87	+	1)
Notidobia ciliaris	3A	3A	1	1	1	1	46	+	
Notonecta glauca	0	0	0	0	0	0	-	+	
Oligochaeta	1	1	1	1	1	1	100		
Orthetrum coerulescens	1	1	1	1	1	1	82	+	
Planorbis sp.	0	0	0	0	0	0	-		
Radix sp.	0	0	0	0	0	0	-		
Simulium latigonium	3A	3A	3A	1	1	1	78	+	
Succinea sp.	3A	3A	1	1	1	1	22		
Sympetrum striolatum	3A	3A	3A	1	1	1	36	+	
Tipula sp.	3A	3A	3A	1	1	1	75	+	
Phylogenetic groups									
Annelida	3A	1	1	1	1	1	31		2)
Crustacea	3A	3A	1	1	1	1	39		
Gastropoda	5A	3A	3A	1	1	1	53		2)
Insecta	5A	5A	3A	3A	1	1	30		
Community assessment									
Analysis type 1	5B	5B	3A	3A	3A	1	32		
Analysis type 2	3A	3A	3A	1	1	1	not analysed		
Analysis type 3	5B	5B	5B	5B	5A	1	17		

1) Significant effect observed shorter than eight weeks. However, due to emergence within eight weeks, no occurrence was observed later. Hence, neither recovery nor longer term effects can be demonstrated. Therefore effect class 5A was selected.
2) Positive effects.

The results presented in Table H.17 can be summarised as follows:

- Several taxa are assigned effect class 0 as no effects are detected because of low abundance or high variance (15 out of 32).

- Few taxa do not show any effect and thus are assigned to effect class 1 for all concentrations (3 out of 32).

- For most of those taxa that show an effect, only a short but pronounced and significant effect was identified (10 out of 15) (effect class 3A). The MDD at the time point of recovery was suitable to identify no effect (below 100 %).

- For few of those taxa that showed an effect, the duration of effects were longer than 8 weeks (3 out of 15) (*Baetis macani*, *Cloeon dipterum*, *Nemoura cinerea*) (effect class 5A/B).

- NOECs at the lowest concentration of 0.1 µg/L were identified for five taxa (*Baetis macani*, *Cloeon dipterum*, *Ischnura elegans*, *Limnephilus lunatus*, *Nemoura cinerea*) (effect class 1). Out of these, two taxa had an effect at the LOEC that did not allow calculation of recovery as individuals emerged (*Baetis macani*, *Nemoura cinerea*).

- Grouping taxa according to phylogenetic attributes identified Annelida, crustaceans and gastropods as relatively insensitive species. In contrast, the group Insecta is rather sensitive with a NOEC (effect class 1) of 0.6 µg/L. Within Insecta, EPT taxa (Ephemeroptera, Plecoptera, Trichoptera) are more sensitive with a pronounced short-term effect (effect class 3A) at 0.6 µg/L, and a NOEC (effect class 1) of 0.1 µg/L.

- Calculating statistical patterns within the dataset (community assessment analysis type 2) revealed short-term effects (effect class 3A) up to 20 µg/L, a NOEC (effect class 1) of 3.2 µg/L.

- Community assessment analysis type 3 enabled detection of long-term effects (effect class 5A) at 0.6 µg/L, and a NOEC (effect class 1) of 0.1 µg/L.

- Overall it can be concluded that population-level analysis and community-level analysis (analysis type 3) in this case leads to the same effect class 1 concentration of 0.1 µg/L and that on the basis of both population-level and community level responses the ecological recovery option cannot be applied (LOEC is an effect class 5A response, while for the ecological recovery option an effect class 3A of the most sensitive population-/community-level endpoint is required).

Following the recommendations in the Aquatic Guidance Document and to address the ecological threshold option, the effect class 1 concentration may be used in the effect assessment by applying an AF of 2 (if the exposure profile in the test system is realistic to worst case when compared with that in the field and the mesocosm system is sensitive, also in terms of species composition). In the experimental streams, the exposure regime is assumed to be realistic to worst case relative to the predicted exposure profiles for edge-of-field surface waters (Figure H.10), although the exposure regime was characterised by a single application only (field exposure is calculated on the basis of two applications at a 30-day interval; Table H.1). Also a sufficient number of sensitive taxa (i.e. more than eight) was present in high enough densities. Based on the effect class 1 (NOEC of 0.1 µg/L) for five taxa, as well as for EPT species and community analysis type 3 species (grouping according to species traits), the **tier 3 ETO-RAC$_{sw;ac}$** might be set at **0.05 µg/L**.

To address the ecological recovery option, the effect class 3A may be used to derive an ERO-RAC by applying an AF of 3–4. However, in this case study no treatment level revealed an overall effect class 3A. At the 0.6 µg/L also effect class 5A was observed for two species. Furthermore, in order to apply the ERO option, vulnerable and sensitive taxa need to be present in the experimental ecosystem study.

In this study, nine vulnerable and sensitive species were present with high enough MDD. As in edge-of-field surface water bodies a high amount of vulnerable taxa are present (sensitive and at the same time long lived), the possibility of recovery cannot be granted assuming a yearly recurring exposure. Hence, for this compound and the specific usage pattern evaluated, the ERO is not applicable since it cannot ensure a sustainable development of sensitive populations. This is in line with the recommendations in the GD, as it is stated that the ecological recovery option (ERO), can accept some population-level effects if ecological recovery takes place within an acceptable time period, which is not the case here. In detail, several endpoints indicate that recovery may not be completed within one year:

- Two taxa (*Baetis macani*, *Nemoura cinerea*) show strong short-term effects with low survival at 0.6 µg/L. These survivors emerged and the potential for recovery of larval stages could not be assessed in the experimental stream study. In such a situation, the guidance document suggests assigning a higher effect class to the treatment level.

- Nine of the individual (non-grouped) taxa reported in Table H.17 had a class IV MDD (10–50 %; small effects can be detected; see Table 31 of Aquatic Guidance Document). Note that the community assessment analysis types 1 and 3 (endpoints constructed by grouping individual taxa on the basis of biological traits) had the lowest MDD (also Class IV) (see Table H.17).

H.3.3 Insecticide I_N chronic effect assessment

H.3.3.1 Tier 1 chronic effect assessment

Chronic toxicity data for standard test species are presented in Table H.18.

Table H.18: Chronic toxicity data for aquatic standard test species and insecticide I_N

	Chronic EC_x/NOEC (µg a.s./L)	**Taxonomy**	**Remark**
Daphnia magna	2 548 (21 d NOEC)	Crustacea; Daphniidae	
Chironomus riparius	1.14 (28 d EC_{10})	Insecta; Chironomidae	
Oncorhynchus mykiss	1200 (98 d NOEC)	Pisces; Salmonidae	
Scenedesmus subspicatus	> 10 000 (72 h E_rC_{50})	Green alga	E_rC_{50} used in effect assessment

The insect *Chironomus* is several orders of magnitude more sensitive than other standard test species.

By applying an AF of 10 to the lowest toxicity value of 1.14 µg/L for *Chironomus riparius* the **tier 1 chronic RAC becomes 0.114 µg/L**.

H.3.3.2 Tier 2 chronic effect assessment on the basis of standard and additional toxicity data

H.3.3.2.1 Tier 2A: Geomean—assessment factor approach

According to the Aquatic Guidance Document the Geomean approach should be performed separately for Crustacea and Insecta.

Besides for *Daphnia magna*, chronic toxicity data for four other crustaceans are available (Table H.**19**).

Table H.19: Chronic toxicity data for standard and additional aquatic test species of Crustacea and insecticide I_N

	Chronic EC_{10}/NOEC µg a.s./L	**Taxonomy; family**
Daphnia magna	2 548 (21 d reproduction)	Crustacea; Daphniidae
Gammarus pulex	7.7 (28 d mortality)	Crustacea; Gammaridae
Hyalella azteca	2.95 (28 d mortality)	Crustacea; Hyalellidae
Neocaridina denticulata	1.71 (28 d mortality)	Crustacea; Atyidae
Asellus aquaticus	1.0 (28 d mortality)	Crustacea; Asellidae

According to the Aquatic Guidance Document, toxicity data based on similar endpoints should be selected when applying the Geomean approach. The geometric mean NOEC/EC_{10} of the toxicity values for crustaceans and the endpoint mortality (n = 4) in the table above is 2.496 µg/L.

For seven insect taxa chronic toxicity data are available (Table H.20).

Table H.20: Chronic toxicity data for standard and additional aquatic test species of Insecta and insecticide I_N

	Chronic NOEC/EC_{10} (µg a.s./L)	**Taxonomy; Family**
Chironomus tentans	2.09 (28 d emergence)	Insecta; Chironomidae (Nematocera)
Chronomus riparius	1.14 (28 d emergence)	Insecta; Chironomidae (Nematocera)
Chaoborus flavicans	1.99 (28 d mortality)	Insecta; Chaoboridae (Nematocera)
Silalis lutaria	2.03 (28 d mortality)	Insecta; Sialidae (Megaloptera)
Plea minutissima	1.28 (28 d mortality)	Insecta; Pleidae (Heteroptera)
Baetis rhodani	0.033 (28 d mortality)	Insecta; Baetidae (Ephemeroptera)
Caenis robusta	0.024 (28 d mortality)	Insecta; Caenidae (Ephemeroptera)

The Geomean NOEC/EC_{10} for the insect toxicity values and all endpoints (emergence and mortality) in the table above is 0.516 µg/L. This can be done as failure to emerge in chironomids in fact can be considered as mortality.

The Geomean for chronic NOEC/EC_{10} values is lowest for insects.

Applying an AF of 10 to the lowest Geomean value (insects) presented in the table above results in a **tier 2A chronic RAC of 0.052 µg/L. Since this value is higher than the lowest chronic toxicity value presented in the Table H.20, the final tier 2A chronic RAC should be ≤ 0.024 µg/L (the lowest chronic NOEC value in Table H.20).**

H.3.3.2.2 Tier 2B: Species sensitivity distribution (SSD) approach

Sufficient chronic toxicity data are available for arthropods but not for crustaceans or insects.

The SSD constructed with chronic toxicity data for aquatic arthropods is presented in Figure H.13. As can be seen the toxicity data do not fit the curve very well and the Anderson–Darling test for normality is rejected at all levels. Therefore, it was decided not to use the SSD approach in the derivation of the chronic RAC.

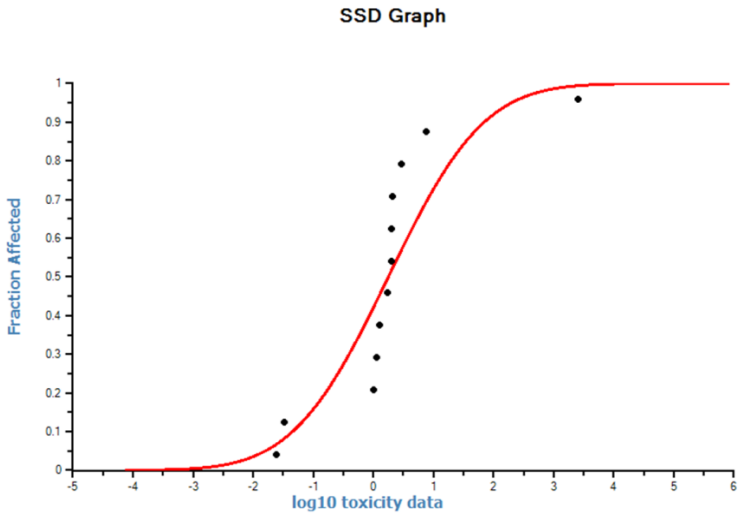


Figure H.13: Species sensitivity distribution (SSD) curve for insecticide I_N constructed with chronic toxicity data from aquatic arthropods (insects and crustaceans)

H.3.3.3 Tier 3: The model ecosystem approach

Under the condition that the exposure regime of the experimental stream study is relatively worst case for the long-term exposure regime in the field, this study might also be used for the long-term risk assessment by comparing the tier 3 RAC (expressed in terms of peak concentrations in the test system) with the $PEC_{sw;max}$ (for results see Table H.21). Since the exposure regime tested in the experimental stream mesocosms is more or less realistic to worst case relative to the predicted exposure profiles for most of the relevant FOCUS scenarios, this procedure was followed.

Following the recommendations in the Aquatic Guidance Document and to address the ecological threshold option the effect class 1 concentration (0.10 µg/L) may be used in the effect assessment by applying an AF of 2 (if the exposure profile in the test system is realistic to worst case when compared with that in the field). Since the exposure regime tested in the experimental ponds is realistic to worst case relative to the predicted exposure profiles for edge-of-field surface waters (Figure H.10), the **tier 3 ETO-$RAC_{sw;ch}$** might be set at **0.05 µg/L** (obligatory to link with $PEC_{sw;max}$). The procedure described above, however, may not be valid for Pond scenarios D5 (see Figure H.10) and D4 (profile comparable to D5), as these exposure profiles are characterised by periods with more or less constant exposure. For these specific cases, the Aquatic Guidance Document offers the possibility to express the responses in the micro-/mesocosm experiment in terms of TWA concentrations measured in these test systems (for further guidance see Aquatic Guidance Document, sections 9.3.5.2–9.3.5.4). When expressing the responses in the pulsed experimental pond study in terms of, for example,. 21-day TWA concentrations (see Table H.18), the tier 3 ETO-$RAC_{sw;ch}$ may be derived by applying an AF of 2 to the 21-day TWA effect class 1 concentration of 0.0.04 µg/L resulting in an **ETO-$RAC_{sw;ch}$ of 0.02 µg/L for D4 and D5 Pond scenarios**. This ETO-$RAC_{sw;ch}$ value for the D4 and D5 pond scenarios may be compared with the $PEC_{sw;max}$ or $PEC_{sw;twa}$ (default seven-day time window), based on expert judgement by considering the criteria mentioned in Aquatic Guidance Document, chapter 4.

H.3.4 Summary of acute and chronic effect assessment insecticide I_N

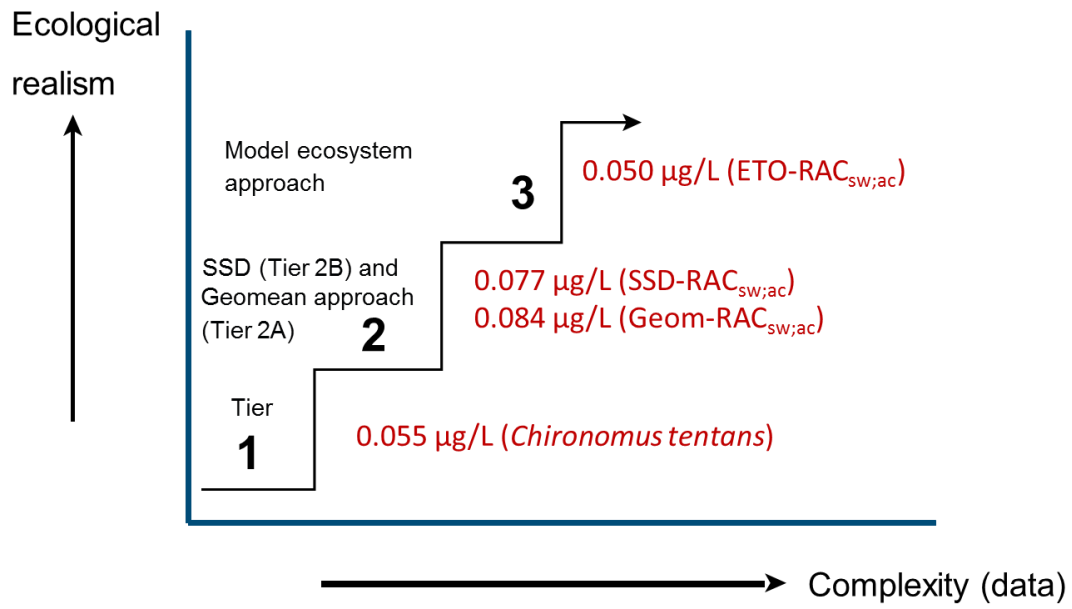


Figure H.14: Schematic presentation of the $RAC_{sw;ac}$ values derived on the basis of different tiers for the insecticide I_N in the acute effect assessment

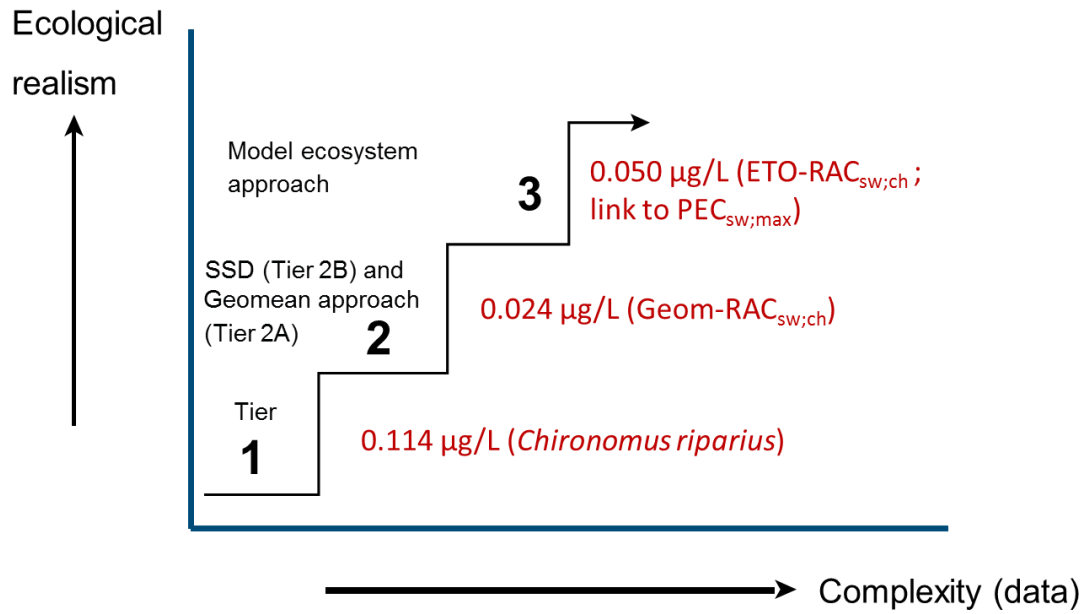


Figure H.15: Schematic presentation of the $RAC_{sw;ch}$ values derived on the basis of different tiers for the insecticide I_N in the chronic effect assessment (except for the D5 Pond scenario)

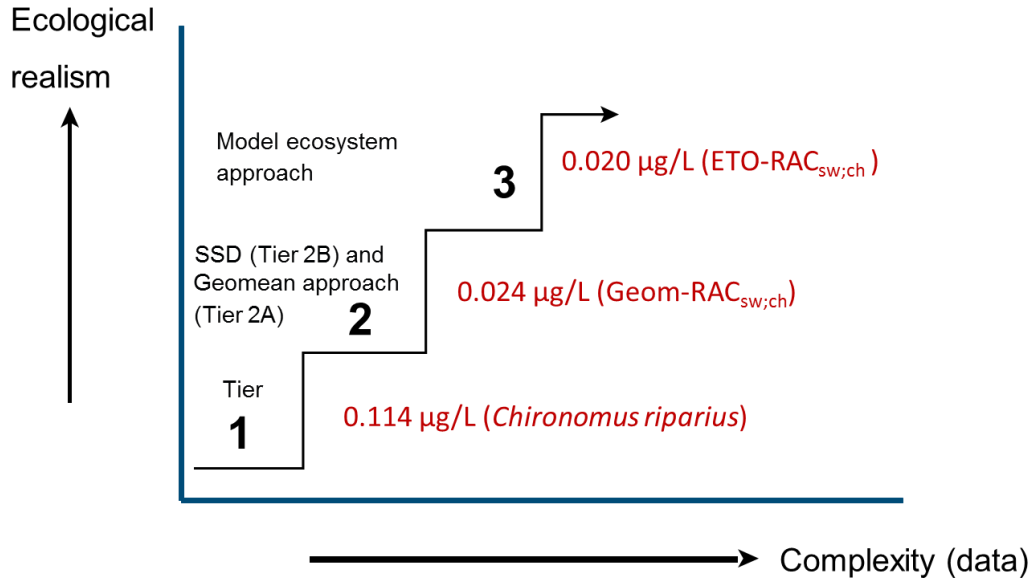


Figure H.16: Schematic presentation of the $RAC_{sw;ch}$ values derived on the basis of different tiers for the insecticide I_N in the chronic effect assessment for the D5 Pond scenarios

H.3.5 Linking exposure to effects in the risk assessment

As final $RAC_{sw;ac}$ and $RAC_{sw;ch}$ values, the tier 3 results are adopted, i.e. the ETO-RAC_{sw} of 0.05 µg/L for all exposure scenarios except the D4 and D5 Pond scenarios. This can be done because, except for the D4 and D5 Pond scenarios, the exposure regime tested in the experimental pond mesocosms is more or less realistic to worst case relative to the predicted exposure profiles for the different FOCUS scenarios. The tier 3 RAC values have to be compared with the $PEC_{sw;max}$ values as calculated for relevant FOCUS scenarios. For the ETO-RAC_{sw} this is done in Table H.21. For the D4 and D5 Pond scenarios the final ETO-$RAC_{sw;ch}$ value selected is 0.02 µg/L.

Table H.21: $PEC_{sw;max}$ values for insecticide I_N for different FOCUS scenarios and different risk-reducing measures (90 % drift-reducing nozzles). The red values indicate that the $PEC_{sw;max}$ is higher than the ETO-RAC_{sw} (0.02 µg/L for D4 and D5 scenarios and 0.05 µg/L for other scenarios) and that under these circumstances the estimated risks are unacceptable. The black and bold value indicate that the $PEC_{sw;max}$ is lower than the ETO-RAC_{sw} and, consequently, that risks are estimated to be low under these circumstances

	$PEC_{sw;max}$ of insecticide I_N (µg/L)	
Scenario	**Step 3**	**Step 4** **90 % drift reduction**
D3 Ditch	7.020	0.703
D4 Pond	0.487	0.487
D4 Stream	7.180	0.793
D5 Pond	0.471	0.311
D5 Stream	7.601	0.803
R1 Pond	0.441	**0.044**
R1 Stream	5.664	1.299
R2 Stream	7.469	0.747
R3 Stream	7.940	0.911
R4 Stream	5.663	1.190

From the results presented in Table H.21 it appears that only for the R1 Pond scenario the risks of exposure to insecticide I_N are estimated to be low if 90 % drift-reducing nozzles are used.

REFERENCES

Arts GHP, Buijse-Bogdan LL, Belgers JDM, Van Rhenen-Kersten CH, Van Wijngaarden RPA, Roessink I, Maund SJ, Van den Brink PJ and Brock TCM, 2006. Ecological impact in ditch mesocosms of simulated spray drift from a crop protection programme for potatoes. Integrated Environmental Assessment and Management, 2, 105–125.

Barnthouse LW, 2004. Quantifying population recovery rates for ecological risk assessment. Environmental Toxicology and Chemistry, 23, 500–508.

Biever RC, Giddings JM, Kiamos M, Annunziato MF, Meyerhoff R and Racke K, 1994. Effects of chlorpyrifos on aquatic microcosms over a range of off-target drift exposure levels. Brighton Crop Protection Conference—Pests and Diseases, 1367–1372.

Brockway DL, Smith PD and Stancil FE, 1984. Fate and effects of atrazine in small aquatic microcosms. Bulletin of Environmental Contamination and Toxicology, 32, 345–353.

Crank J, 1967. The mathematics of diffusion. Oxford University Press, Oxford, UK.

Cuppen JGM, Van den Brink PJ, Camps E, Uil KF and Brock TCM, 2000. Impact of the fungicide carbendazim in freshwater microcosms. I. Water quality, breakdown of particulate organic matter and responses of macroinvertebrates. Aquatic Toxicology, 48, 233–250.

Daam MA, Crum SJH, Van den Brink PJ and Nogueira AJA, 2008. Fate and effects of the insecticide chlorpyrifos in outdoor plankton-dominated microcosms in Thailand. Environmental Toxicology and Chemistry, 27, 2530–2538.

Daam MA, Satapornvanit K, Van den Brink PJ and Nogueira AJA, 2009. Sensitivity of macroinvertebrates to carbendazim under semi-field conditions in Thailand: implications for the use of temperate toxicity data in a tropical risk assessment of fungicides. Chemosphere, 74, 1187–1194.

DeNoyelles F Jr, Dewey SL, Huggins DG and Kettle WD, 1994. Aquatic mesocosms in ecological effects testing: detecting direct and indirect effects of pesticides. In: Aquatic mesocosm studies in ecological risk assessment. Eds Graney RL, Kennedy JH and Rodgers JH Jr. Lewis Publishers, Boca Raton, FL, , USA, 577–603.

Detenbeck NE, Hermanutz R, Allen K and Swift MC, 1996. Fate and effects of the herbicide atrazine in flow-through wetland mesocosms. Environmental Toxicology and Chemistry, 15, 937–946.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2010. Scientific opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the Guidance Documents on Aquatic and Terrestrial Ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002). EFSA Journal 2010;8(10):1821, 55 pp. doi:10.2903/j.efsa.2010.1821.

Fairchild JF, La Point TW and Schwartz TR, 1994. Effects of herbicide and insecticide mixture in aquatic mesocosms. Archives of Environmental Contamination and Toxicology, 27, 527–533.

Farmer D, Hill IR and Maund SJ, 1995. A comparison of the fate and effects of two pyrethroid insecticides (lambda-cyhalothrin and cypermethrin) in pond mesocosms. Ecotoxicology, 4, 219–244.

FOCUS, 2000. FOCUS groundwater scenarios in the EU review of active substances. Report of the FOCUS Groundwater Scenarios Workgroup, EC Document Reference SANCO/321/2000 rev. 2, 202 pp.

FOCUS, 2001. FOCUS Surface Water Scenarios in the EU Evaluation Process under 91/414/EEC. Report of the FOCUS Working Group on Surface Water Scenarios, EC Document Reference SANCO/4802/2001 rev. 2, 245 pp.

FOCUS, 2007a. Landscape and mitigation factors in aquatic risk assessment. Volume 1. Extended Summary and Recommendations. Report of the FOCUS Working Group on Landscape and

Mitigation Factors in Ecological Risk Assessment, EC Document Reference SANCO/10422/2005 v2.0, 169 pp.

FOCUS, 2007b. Landscape And Mitigation Factors In Aquatic Risk Assessment. Volume 2. Detailed Technical Reviews. Report of the FOCUS Working Group on Landscape and Mitigation Factors in Ecological Risk Assessment, EC Document Reference, SANCO/10422/2005 v2.0, 436 pp.

Fryday S and Thompson H, 2012. Toxicity of pesticides to aquatic and terrestrial life stages of amphibians and occurrence, habitat use and exposure of amphibian species in agricultural environments. Supporting Publications 2012:EN-343. [348 pp.]. Available online: http://www.efsa.europa.eu/en/supporting/doc/343e.pdf

Gruessner B and Watzin MC, 1996. Response of aquatic communities from a Vermont stream to environmentally realistic atrazine exposure in laboratory microcosms. Environmental Toxicology and Chemistry, 15, 410–419.

Guasch H and Sabater S, 1998. Light history influences the sensitivity to atrazine in periphytic algae. Journal of Phycology, 34, 233–241.

Heckman LH and Friberg N, 2005. Macroinvertebrate community response to pulse exposure with the insecticide lambda-cyhalothrin using in-stream mesocosms. Environmental Toxicology and Chemistry, 24, 582–590.

Hill IR, Runnalls JK, Kennedy JH and Ekoniak P, 1994. Lambda-cyhalothrin: a mesocosm study of its effects on aquatic organisms. In: Aquatic mesocosm studies in ecological risk assessment. Eds Graney RL, Kennedy JH and Rodgers JH Jr. .Special Publication of the Society of Environmental Toxicity and Chemistry. Lewis Publishers, Michigan, USA, 403–467.

Johnson BT, 1986. Potential impact of selected agricultural chemical contaminants on a northern prairie wetland: a microcosm evaluation. Environmental Toxicology and Chemistry, 5, 473–485.

Jüttner I, Peither A, Lay JP, Kettrup A and Ormerod SJ, 1995. An outdoor mesocosm study to assess ecotoxicological effects of atrazine on a natural plankton community. Archives of Environmental Contamination and Toxicology, 29, 435–441.

Kosinski RJ, 1984. The effect of terrestrial herbicides on the community structure of stream periphyton. Environmental Pollution Series A, 36, 165–189.

Kosinski RJ and Merkle MG, 1984. The effect of four terrestrial herbicides on the productivity of artificial stream algal communities. Journal of Environmental Quality, 13, 75–82.

Krieger KA, Baker DB, Kramer JW, 1988. Effects of herbicides on stream Aufwuchs productivity and nutrient uptake. Archives of Environmental Contamination and Toxicology, 17, 299–306.

Lee AFS and Gurland J, 1975. Size and power of tests for equality of means of two normal populations with unequal variances. Journal of the American Statistical Association, 70, 933–941.

López-Mancisidor P, Carbonell G, Marina A, Fernández C and Tarazona JV, 2007. Zooplankton community responses to chlorpyrifos in mesocosms under Mediterranean conditions. Ecotoxicology and Environmental Safety, 71, 16–25.

López-Mancisidor P, Carbonell G, Fernández C and Tarazona JV, 2008. Ecological impact of repeated applications of chlorpyrifos on zooplankton community in mesocosms under Mediterranean conditions. Ecotoxicology, 17, 811–825.

Lozano SJ, O'Halloran SL, Sargent KW and Brazner JC, 1992. Effects of esfenvalerate on aquatic organisms in littoral enclosures. Environmental Toxicology and Chemistry, 11, 35–47.

Maltby L, Arnold D, Arts G, Davies J, Heimbach F, Pickl C and Poulsen V (Eds), 2010. Aquatic macrophyte risk assessment for pesticides. SETAC Press & CRC Press, Boca Raton, FL, USA, 140 pp.

Muños I, Real M, Guasch H, Navarro E and Sabater S, 2001. Effects of atrazine on periphyton under grazing pressure. Aquatic Toxicology, 55, 239–249.

Nyström B, Paulsson M, Almgren K and Blank H, 2000. Evaluation of the capacity for development of atrazine tolerance in periphyton from a Swedish freshwater site as determined by inhibition of photosynthesis and sulfolipid synthesis. Environmental Toxicology and Chemistry, 19, 1324–1331.

Pratt JR, Bowers NJ, Niederlehner BR and Cairns J Jr, 1988. Effects of atrazine on freshwater microbial communities. Archives of Environmental Contamination and Toxicology, 17, 449–457.

Pusey BJ, Arthington AH and McClean J, 1994. The effects of a pulsed application of chlorpyrifos on macroinvertebrate communities in an outdoor artificial stream system. Ecotoxicology and Environmental Safety, 27, 221–250.

Roessink I, Arts GHP, Belgers JDM, Bransen F, Maund SJ and Brock TCM, 2005. Effects of lambda-cyhalothrin in two ditch microcosm systems of different trophic status. Environmental Toxicology and Chemistry, 24, 1684–1696.

Seguin F, Leboulanger C, Rimet F, Druart JC and Berard A, 2001. Effects of atrazine and nicosulfuron on phytoplankton in systems of increasing complexity. Archives of Environmental Contamination and Toxicology, 40, 198–208.

Siefert RE, Lozano SJ, Brazner JC and Knuth ML, 1989. Littoral enclosures for aquatic field testing of pesticides: effects of chlorpyrifos on a natural system. Miscellaneous Publication Series Entomological Society of America, 75, 57–73.

Slijkerman DME, Baird DJ, Conrad A, Jak RG and Van Straalen NM, 2004. Assessing structural and functional plankton responses to carbendazim toxicity. Environmental Toxicology and Chemistry, 23, 455–462.

Stampfli NC, Knillmann S, Liess M and Beketov MA, 2011. Environmental context determines community sensitivity of freshwater zooplankton to pesticides. Aquatic Toxicology, 104, 116–124.

Stay FS, Katko A, Rohm CM, Fix MA and Larsen DP, 1989. The effects of atrazine on microcosms developed from four natural plankton communities. Archives of Environmental Contamination and Toxicology, 18, 866–875.

Van den Brink PJ, Van Donk E, Gylstra R, Crum SJH and Brock TCM, 1995. Effects of chronic low concentrations of the pesticides chlorpyrifos and atrazine in indoor freshwater microcosms. Chemosphere, 31, 3181–3200.

Van den Brink PJ, Van Wijngaarden RPA, Lucassen WGH, Brock TCM and Leeuwangh P, 1996. Effects of the insecticide Durban 4®E (a.i. chlorpyrifos) in outdoor experimental ditches: II. Community responses and recovery. Environmental Toxicology and Chemistry, 15, 1143–1153.

Van den Brink PJ, Hattink J, Bransen F, Van Donk E and Brock TCM, 2000. Impact of the fungicide carbendazim in freshwater microcosms. II. Zooplankton, primary producers and final conclusions. Aquatic Toxicology, 48, 251–264.

Van den Brink PJ, Hartgers EM, Gylstra R, Bransen F and Brock TCM, 2002. The effects of a mixture of two insecticides on freshwater microcosms. II. Water quality, responses of zooplankton, phytoplankton and periphyton and ecological risk assessment. Ecotoxicology, 11, 181–197.

Van Wijngaarden RPA, Cuppen JGM, Arts GHP, Crum SHJ, Van den Hoorn MW, Van den Brink PJ and Brock TCM, 2004. Aquatic risk assessment of a realistic exposure to pesticides used in bulb crops: a microcosm study. Environmental Toxicology and Chemistry, 23, 1479–1498.

Van Wijngaarden RPA, Brock TCM and Douglas MT, 2005. Effects of chlorpyrifos in freshwater model ecosystems: the influence of experimental conditions on ecotoxicological thresholds. Pest Management Science, 61, 923–935.

Van Wijngaarden RPA, Brock TCM, Van den Brink PJ, Gylstra R and Maund SJ, 2006. Ecological effects of spring and late summer applications of lambda-cyhalothrin in freshwater microcosms. Archives of Environmental Contamination and Toxicology, 50, 220–239.

Verdonck FAM, Aldenberg T, Jaworska J and Vanrolleghem PA, 2003. Limitations of current risk characterization methods in probabilistic environmental risk assessment. Environmental Toxicology and Chemistry, 22, 2209–2213.

Webber EC, Deutch WG, Bayne DR and Seesock WC, 1992. Ecosystem-level testing of synthetic pyrethroid insecticide in aquatic mesocosms. Environmental Toxicology and Chemistry, 11, 87–105.