SCIENTIFIC OPINION

Scientific Opinion addressing the state of the science on risk assessment of plant protection products for non-target terrestrial plants[1]

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

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

Following a request from the European Food Safety Authority, the Panel on Plant Protection Products and their Residues developed an opinion on the science to support the development of a risk assessment scheme of plant (crop) protection products on non-target terrestrial plants (NTTPs). This scientific opinion is largely a literature review on the most up-to-date knowledge of factors influencing phytotoxicity testing and risk assessment of NTTPs. Specific protection goals (SPGs) were defined for off-field, in-field and endangered species. SPGs are closely linked to ecosystem services and functions, and include maintaining provision of water regulation, food web support, aesthetic values, genetic resources and biodiversity. Gaps were identified in standard guidelines currently used in lower tier testing (tier I/II). In these guidelines, tests are conducted at the seedling/juvenile stage using mostly annual crops, and effects are recorded at the juvenile/vegetative stage under greenhouse conditions with plants grown individually or in monoculture. Endpoints measured do not include the overall effect on the whole life cycle (germinating seeds, seedling, juvenile stages, flowering, and seed production and germinability). It is also noted that it is unknown whether the following groups of organisms are covered by the plant risk assessment as it is carried out now: ferns, mosses, liverworts, hornworts, horsetails, lichens or woody species. In terms of exposure, droplet drift is considered to be the most important factor for off-field emissions to non-target areas. Models are available to calculate loadings from spray drift. Higher tier assessment is not required if the risk based on the tier II level can be managed by risk mitigation measures. When required, higher tier tests should be conducted under more realistic conditions. They may include additional laboratory/greenhouse tests (e.g. to measure reproductive endpoints or species interactions), microcosms or field experiments with experimentally or already established species. Other issues were considered, including exposure to mixtures, adjuvants, co-formulants and metabolites. Recommendations for the improvement of current guidelines and the elaboration of new guidelines and risk assessment schemes are provided.

[1] On request from EFSA, Question No EFSA-Q-2011-00982, adopted on 10 July 2014.

[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 and Ton Van Der Linden. Correspondence: pesticides.ppr@efsa.europa.eu

[3] Acknowledgement: The Panel wishes to thank the members of the Working Group on Non-target terrestrial plant RA: Céline Boutin, Michael Klein, Robert Luttik, Mark Rees, Carmen Schweikert, Walter Steurbaut and Beate Strandberg, and EFSA staff: Stephanie Bopp, Franz Streissl, Laura Villamar Bouza 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), 2014. Scientific Opinion addressing the state of the science on risk assessment of plant protection products for non-target terrestrial plants. EFSA Journal 2014;12(7):3800, 163 pp. doi:10.2903/j.efsa.2014.3800

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

KEY WORDS

non-target terrestrial plants, phytotoxicity testing, species selection, droplet drift exposure, whole life cycle, Specific Protection Goal, risk mitigation measures

SUMMARY

The European Food Safety Authority (EFSA) asked the Panel on Plant Protection Products and their Residues (PPR) to develop and update the guidance documents on terrestrial ecotoxicology under mandate M-2009-0002. The assessment of effects on biodiversity is not explicitly addressed under the existing guidance documents and appropriate risk assessment methodology needs to be developed. As such, expertise was needed in the different areas of terrestrial ecotoxicology, including non-target terrestrial plants (NTTPs). This scientific opinion, which is largely a literature review, has been written as a precursor to the guidance document on NTTPs.

Non-target terrestrial plants were defined as all plants growing outside fields, and those growing within fields that are not the intended pesticide target. The temporal and spatial boundaries of the off-field areas were defined. The overall protection goal for higher terrestrial plants is to maintain the biodiversity of plant species in the agricultural area, including both the above- and belowground (seed bank) diversity, and is linked to ecosystem services. Specific protection goals were defined: (1) for off-field NTTPs as key drivers for nutrient cycling, water regulation, food web support, aesthetic values and genetic resources (biodiversity); (2) for in-field NTTPs as key drivers for food web support (primary production, provision of habitat and food for other non-target organisms, e.g. arthropods, birds), aesthetic values and genetic resources; and (3) for endangered species including rare arable weeds.[4] Plants are primary producers and recognised as the foundation of terrestrial ecosystems. Plant recovery from the application of plant protection products (PPPs),[5] effects on reproduction (as well as effects on vegetative parts) and consequences to other trophic levels are considered. Available risk mitigation options for the in- and off-field risks are presented.

Two organisations have developed guidelines for testing phytotoxicity to crops and NTTPs: the Organisation for Economic Co-operation and Development (OECD) and the United States Environmental Protection Agency (US EPA). In all cases, tests are conducted at the seedling/juvenile stage and effects recorded at the vegetative stage under greenhouse conditions with plants grown individually or in monoculture. Tier I encompasses testing at the maximum label rate and tier II aims to obtain dose–response curves. Although a list of 32 crop and 52 non-crop species historically used in phytotoxicity testing is identified in guidelines (OECD and US EPA), tests provided by registrants are mostly on crop species. Additional efficacy data generated to establish action on weeds and following or neighbouring crops are also provided in the dossier for registration. These data are valuable as they offer easily accessible information on the sensitivity of a wide range of plant species, many of which are important to wildlife. In addition, important information on the range of sensitivity among crop varieties (not provided in regulatory testing) can be obtained with these efficacy data. In all cases, tests are conducted at the seedling/juvenile stage and effects recorded at the vegetative stage under greenhouse conditions with plants grown individually or in monoculture.

Species representation in testing of PPPs in NTTPs was considered incomplete in the current guidelines. Tests provided for the authorisation of PPPs are mostly conducted with annual crop species and only one variety (usually not defined) per crop. Ecologically relevant species (NTTPs) are not frequently tested. Experiments have demonstrated that crops may or may not be suitable surrogates for wild species (herbaceous and woody). No tests are conducted on ferns, mosses, liverworts, hornworts, horsetails and lichens, and limited data indicate that these taxa are quite sensitive. Perennial or woody species are also neglected. Annual and perennial species do not consistently differ in their toxicological sensitivity to herbicides. Tests are conducted on crops sprayed at the juvenile stage (two- to four-leaf stages), with effects measured 21–28 days after spraying. Research conducted on woody and herbaceous species showed that they may be very sensitive when sprayed at the reproductive stage, and this is not tested. Many non-crop species are deemed suitable for phytotoxicity testing as they germinate readily and grow uniformly under greenhouse conditions with minimum requirements.

[4] The term weeds is used in the current document for undesired plants and for non-target plants in the in-field area.

[5] The term plant protection product is used for substances intended to protect crops or crop products against harmful organisms, to influence the life processes of crops or to destroy undesired plants, i.e. weeds, or parts of these. Therefore, the term plant protection product as used in this opinion also includes herbicides.

Endpoints recommended in the standard guidelines (OECD and US EPA) were reviewed. Endpoints usually measured include visual assessment, aboveground biomass and height. When PPPs are applied, NTTPs can be at the seedling/juvenile or reproductive stages, and effects can be observed on either the vegetative or the reproductive parts. Effects can be seen immediately or be delayed. In current guidelines, plants are tested only at the seedling/juvenile stage. It was found in the literature that plants at the reproductive stage were affected differently, and that the reproductive endpoints (e.g. flower and seed production) were, in general, more sensitive than vegetative measures. It was concluded that effects on the whole life cycle (germinating seeds, seedling, juvenile stages, flowering, and seed production and germinability) have to be considered to properly assess the effects on natural NTTP populations and facilitate the protection of biodiversity. It is also acknowledged that not just herbicides can affect non-target plants. There may be adverse effects, for example from exposure to fungicides via impacts on mycorrhiza.

Effect endpoints for non-target plants are traditionally expressed per area rather than as concentrations. Consequently, exposure should also be expressed in the same unit (e.g. grams of active substance per hectare). Droplet drift is currently considered to be the most important factor for off-field emissions to non-target areas, but vapour and dust drift can also contribute to off-field deposition. Exposure models to calculate loadings from droplet and vapour drift are available and could be used to assess exposure for non-target plants. It should be noted, however, that spray drift values in field crops originating from recent research were considerably higher than the values that are currently used in the exposure assessment at EU level. No models are available for dust drift. Surface run-off may contribute to the contamination of non-target terrestrial ecosystems in the neighbourhood of agricultural areas. Models to estimate run-off losses are available and used for the assessment of the aquatic environment. Other emission routes such as leaching and drainage are generally not considered as direct emission routes in the terrestrial compartment.

Plants as primary producers are the foundation of terrestrial ecosystems, and measures of the ecosystem services and functions they provide should be considered in higher tier testing. The risk to NTTPs of non-herbicide substances are usually sufficiently addressed with tier I tests, whereas, for herbicides/growth regulators and PPPs with herbicidal or growth-regulating activity, the risk needs to be addressed with tier II studies. Higher tier assessment is not required if the risk based on the tier II level can be managed by risk mitigation measures. In current guidelines, no recommendations for higher tier testing are proposed. Higher tier tests should be conducted under more realistic conditions than Tier I/II tests. These may include additional laboratory/greenhouse tests (e.g. to measure reproductive endpoints or species interactions), microcosms or field experiments with planted or already established species. Discrepancies between single- and multi-species tests, as well as reproducibility, need to be assessed. Several available models are presented for studies of the impact of plant interactions/plant competition on plant population and community dynamics based on observed changes in density, biomass or plant cover.

Other issues are considered in this opinion, including exposure to mixtures, adjuvants, co-formulants, metabolites and multiple application factors. It was found that plants and other wildlife are usually exposed to mixtures of compounds in tank mixtures as well as following sequential applications to crops. Effects of mixtures can be evaluated by conducting supplementary tests or by modelling approaches. During spray, NTTPs are exposed not only to the active substance but also to adjuvants and co-formulants that are part of the pesticide formulation and that determine the efficacy of herbicidal pesticides. Consequently, PPP formulations, and if appropriate actual tank mixtures, were deemed more appropriate for terrestrial plant testing than using an active substance alone, as also recommended in the OECD and US EPA guidelines. Metabolites or degradation products from active substances may be as toxic as the parent compounds and it is recognised that they may need to be tested separately.

There is an ongoing debate regarding the optimum number of species that should be tested in phytotoxicity trials and in support of risk assessment schemes. In current guidelines, 6–10 species are recommended. In lieu of testing a large number of species, using a plant trait-based approach (species

with different physiological, morphological or ecological traits) may be a promising avenue for plant species selection in phytotoxicity testing and the ensuing ecological risk assessment. The use of statistical techniques (such as species-sensitivity distribution model) may also be a way to address the difficulty related to the appropriate number of test species. The implications on population and community levels require further extrapolations.

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

The PPR panel is tasked with the update of the Guidance Document on Terrestrial Ecotoxicology under mandate M-2009-0002. The Guidance Documents that are still in place were developed under Directive 91/414/EEC.

A public consultation on the existing Guidance Documents was held by EFSA in 2008 in order to collect input for the revision of the aquatic and terrestrial Guidance Documents. The following points were most often mentioned in the comments for updating the Guidance Documents:

- Considerations of the revision of Annexes II and III of Directive 91/414/EEC,
- Consideration of the new Regulation (EC) 1107/2009[6]
- Harmonisation with other directives and regulations (biocides, REACH)
- Clearly defined protection goals
- Multiple exposure
- Inclusion of additional species in the risk assessment (e.g. amphibians, reptiles, bats, molluscs, ferns, mosses, lichens, butterflies, grasshoppers and moths)
- More guidance on statistical analysis
- Preference of EC_x over NOEC values in the risk assessment
- To consider all available information from workshops (EUFRAM, ESCORT, PERAS and other SETAC workshops)
- Endocrine disruption
- Consideration of all routes of exposure
- Bee risk assessment
- Non-target arthropods risk assessment
- Soil organism risk assessment

The comments received in the stakeholder consultation will be consulted on again during the revision of the Guidance document.

A survey on the needs and priorities regarding Guidance Documents was conducted among Member States Authorities and a final list was compiled in the Pesticide Steering Committee meeting in November and December 2010.

The following topics were indicated as priorities for the update of the terrestrial Guidance Document:

- Assessment of impacts on non-target organisms including the ongoing behaviour
- Impact on biodiversity

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

- Impact on the ecosystem
- Effects on bees
- Effects on amphibians and reptiles
- Linking exposure to effects and ecological recovery
- The use of field studies in the risk assessment and guidance for interpretation of field studies
- Revision of non-target arthropod risk assessment (ESCORT II)
- Guidance for risk assessment in greenhouses
- Definitions of environmental hazard criteria (POP, PBT, vPvB) which will serve as a cut-off criteria according to the new regulation. Guidance on what studies, test conditions and endpoints should be used in determining whether the cut-off values have or have not been met. The Commission will consider the respective competencies of institutions regarding this topic and will check whether it takes the lead in this area.
- Definition of hazard criteria in relation to endocrine disruption and guidance on what studies, test conditions and endpoints should be used in determining whether the cut-off values have or have not been met. The Commission has the lead in developing these criteria. It is expected that EFSA will be consulted by the Commission on the final report in October 2011. The outcome of these activities should be incorporated in the Guidance Documents.

Generic questions that arose during the peer-review expert meetings should also be taken into consideration in the update of the guidance document. A compilation of general reports was provided by the pesticides unit. One of the points mentioned was that more detailed guidance is needed for the risk assessment of non-target plants (e.g. sensitivity of test species, use of species sensitivity distributions, exposure estimates).

Regulation (EC) 1107/2009 states that the use of plant protection products should have no unacceptable effects on the environment. The regulation lists in particular effects on non-target species, including on their ongoing behaviour and impact on biodiversity and the ecosystem.

The assessment of effects on ongoing behaviour and biodiversity are not explicitly addressed under the existing Guidance Documents and appropriate risk assessment methodology needs to be developed.

The expertise needed in the different areas of terrestrial ecotoxicology ranges from in-soil biology, non-target arthropods, bees and other pollinating insects, terrestrial non-target plants, amphibians and reptiles and modelling approaches in the risk assessment.

This justifies the need to split the activity in several separate areas due to the complexity of the task and in order to make most efficient use of resources.

A separate question was received from the European Commission to develop a Guidance Document on the Risk Assessment of Plant Protection Products for bees and to deliver an opinion on the science behind the risk assessment Guidance. This question will be dealt with under mandate M-2011-0185.

TERMS OF REFERENCE AS PROVIDED BY EFSA

EFSA tasks the Pesticides Unit and the PPR Panel on the following activities taking into consideration Regulation (EC) 1107/2009, stakeholder comments and the recommendations and priorities identified by Member States:

Development of Guidance on risk assessment for non-target terrestrial plants, with the following deliverables:

- Opinion addressing the state of the science to be delivered by the PPR Panel by July 2014
- Guidance of EFSA to be delivered by September 2015
- Public consultation on the draft Guidance of EFSA

This PPR Panel Opinion is the first deliverable under the respective Terms of Reference above.

ASSESSMENT

1. Introduction

Some plant protection products (PPPs) (i.e. herbicides and plant growth regulators) are deliberately released into the agri-environment to control undesired plants. However, these substances and other PPPs may unintentionally affect non-target terrestrial plants (NTTPs) including rare arable weeds. NTTP species mainly occur in the off-field area, although they may also be present inside the fields, that is "in- field" NTTP (see section 2.3). This section provides an introduction to the regulatory requirements and objectives of this opinion.

1.1. Legislative background

Active substances (a.s.) used in PPPs are authorised in the EU under Regulation (EC) No 1107/2009.[7] The regulation requires that "substances or products produced or placed on the market do not have any harmful effect on human or animal health or any unacceptable effects on the environment". With respect to the environment, this includes in particular considerations of the impact on non-target species, including on the ongoing behaviour of those species, and the impact on biodiversity and the ecosystem.

New Commission regulations laying down the data requirements for the dossier to be submitted for the approval of active substances contained in PPPs (Commission Regulation (EU) No 283/2013[8] and 284/2013[9]) were published in 2013. These documents provide information on the core data required for the authorisation of PPPs. Furthermore, as a general requirement for substance approval, it is stated in Commission Regulation (EU) No 283/2013 that "the potential impact of the active substance on biodiversity and the ecosystem, including potential indirect effects via alteration of the food web, shall be considered".

For terrestrial plants, in a first step, screening data shall be required for all active substances, which will allow identification of substances that exhibit herbicidal or plant growth regulator activities, unless it is already known that an active substance exhibits herbicidal or plant growth regulatory activity. Screening data should be provided for at least six plant species from six different families, including mono- and dicotyledons. The tested concentrations and rates shall be equal to or higher than the maximum recommended application rate and either at a rate appropriate to simulate use patterns under field conditions (with testing conducted after the final treatment) or at a rate applied directly that takes into account the accumulation of residues following multiple applications of the PPP.

Furthermore, a summary of available data from tests used to assess biological activity and dose range finding studies, whether positive or negative, shall be provided. This information should also be assessed regarding the potential impact on non-target plant species. These data shall be supplemented by further information, in summary form, on the effects on plants observed during field testing, for efficacy, residues, and environmental fate and ecotoxicological effects.

For substances where screening studies do not cover the required range of species and concentrations or for those exhibiting herbicidal or plant growth regulatory activity, further plant testing shall be carried out. Vegetative vigour and seedling emergence concentration/response tests shall be provided for at least six species representing families for which herbicidal/plant growth regulatory action has

[7] 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, p. 1–50.

[8] 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.

[9] 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.

been found. Where, from the mode of action, it can be clearly established that either seedling emergence or vegetative vigour is affected, only the relevant study shall be conducted.

Data are not required where exposure is negligible, for example in the case of rodenticides, in the case of the active substance used for wound protection or seed treatment, or in the case of the active substances used on stored products or in glasshouses where exposure is precluded. However, where the available information on active substance indicates that unacceptable risk to terrestrial plants cannot be excluded, then such waivers might not be feasible. This may apply for example to highly volatile herbicidal substances and plant growth regulators used in glasshouses.

1.2. The process adopted to revise the terrestrial guidance document

In view of the publication and entry into force of the new Regulation (EC) No 1107/2009 and the revised data requirements for chemical PPPs, as well as new scientific findings, the Panel on Plant Protection Products and their Residues (PPR) was asked to revise the Guidance Document on Terrestrial Ecotoxicology (EC, 2002a). It was decided to split the task and to address individually the risk assessment for separate organism groups, i.e. in-soil organisms, non-target arthropods, amphibians and reptiles, and NTTPs. For each of the organism groups, the PPR Panel will first summarise in a scientific opinion the science behind the respective risk assessment and, in a second step, EFSA will develop practical guidance on how to perform the risk assessment. The risk assessment guidance will mainly focus on the assessment of chemical PPPs. For microbial formulations, specific risk assessment methodologies may need to be developed.

2. Defining specific protection goals

2.1. Protection goals for non-target plant species

The prerequisite for the approval of a PPP according to chapter 1.1 of Regulation (EC) No 1107/2009 is the absence of unacceptable effects on the environment including biodiversity. This requirement in risk regulation of PPPs is reflected in the introduction of the new data requirements, Regulation (EC) No 283/2013, on the assessment of impacts on non-target species (see paragraph 1.11 points d, e and f):

- evaluate the impact on biodiversity and the ecosystem;
- identify non-target species and populations for which hazards arise because of potential exposure;
- permit an evaluation of short- and long-term risks for non-target species, populations, communities and processes.

In addition, under point 5 of the introduction to chapter 8 of (EC) No 283/2013, it is mentioned that "the potential impact of the active substance on biodiversity and the ecosystem, including potential indirect effects via alteration of the food web, shall be considered".

In an agricultural landscape, NTTPs can be defined as all terrestrial plants affected by pesticides, although they are not the intended target of the pesticides. NTTPs can be found within fields (in-field NTTPs) and in non-target areas outside fields (off-field NTTPs). Because herbicides are intentionally used for affecting weeds in the field, it is important to define specific protection goals (SPGs) for in-field plant species as well, since they are an important component in the provision of some ecosystem services. Note that other PPPs with herbicidal activity will also affect weeds.

Although it is currently obvious that the risk of off-field NTTPs should be addressed, it is less clear if the risk of in-field NTTPs should be addressed. However, in the case of herbicides, it is clear that in-field non-crop species are strongly affected and that this could lead to major ecological impacts. For example, in-field NTTPs provide very important ecosystem services in terms of food web support (food and habitat provision). Their reduced abundance in fields strongly impacts farmland species

from higher trophic levels, especially birds. Additionally, wild plants in non-target areas outside fields (off-field NTTPs) are important for the maintenance of biodiversity and the food web. However, these off-field NTTPs are facing an increasing fragmentation of their native habitats, and their presence in fields appears to be of most importance for restoring their metapopulation selection level and maintaining biodiversity at the landscape level. Finally, there are non-crop plant species that only/mainly occur in the in-field area (see section 2.4). These rare arable weeds often grow within fields and may need protection under some circumstances.

For these various reasons, it is important to define SPGs for in-field NTTP species as well as to address their risk in fields, although, in the case of herbicide use, it is problematic to define non-target species among plant species present in the field. Furthermore, it is relevant to point out that the new version of the data requirements Commission Regulations (EU) No 283/2013 and No 284/2013 of 1 March 2013, which came into force after the publication of the SPG opinion (EFSA PPR Panel, 2010), no longer provides the following definition: "non-target plants are plants outside the cropped area". Therefore, there is a legal requirement to explicitly also consider non-target plants in the field as well as in the off-field area.

The protection goals have to allow for the facts that: (1) according to Regulation (EC) 1107/2009 the risk assessment is based on one single compound or the intended use of a PPP, and (2) "Member States shall ensure that the use of PPPs does not have any long-term repercussions for the abundance and diversity of non-target species" (Regulation (EU) No 546/2011). In this context, it should be noted that simultaneous (tank mixtures) as well as sequential multiple exposure (see section 8.1) to PPPs could aggravate the effects on non-target plants in both in- and off-field areas. As a consequence, to meet protection goals, this aspect of multiple exposure has to be taken into account in the risk assessment.

2.2. Ecosystem services relevant for agricultural areas and driven by non-target terrestrial plants

As described in the opinion on SPGs (EFSA PPR Panel, 2010), the concept chosen to derive SPGs was based on the "ecosystem service" framework, as presented in the Millennium Ecosystem Assessment (2005). An overview of the ecosystem services is listed in the opinion on SPGs (EFSA PPR Panel, 2010; see Table 1).

Table 1: Ecosystem services in different spatial areas, their importance in these areas (+ small; ++ intermediate; +++ large) and the potential impact of pesticides (owing to normal agricultural use) on them

Ecosystem service		In crop	Off crop		Strongly impacted by pesticides (direct or indirect effects via trophic interactions)	Potentially impacted taxa
			Small edge of field margins	More remote terrestrial areas (e.g. hedges, natural areas)		
Provisioning services	Food	+++	+	++	Yes	Crop species, cattle, small game and other consumable vertebrates, fungi and berries (wild fruits), consumable fish, crayfish, molluscs, algae
	Fibre and fuel	+++	+	++	Yes	Crop plants (fibres/biofuel), trees (wood/biofuel), emergent macrophytes (thatched roofs), aquatic primary producers and peat (biofuel)
	Energy (hydroelectric and cooling water)				No	Fouling organisms
	Transport (waterways, e.g. boat traffic)				No	Fouling organisms
	Genetic resources/biodiversity	++	++	+++	Yes	All species
	Biochemical/natural medicines	++	+	++	No	Organisms used for medicinal or personal care products
	Ornamental resources	++	++	++	Yes	Ornamental species and landscape elements
	Fresh water	+	++	+++	Yes	Microorganisms, algae, etc.
Regulatory services (beneficial regulations)	Pollination	+++	+++	+++	Yes	Bees and other pollinator species (particularly insects)

Ecosystem service		In crop	Off crop		Strongly impacted by pesticides (direct or indirect effects via trophic interactions)	Potentially impacted taxa
			Small edge of field margins	More remote terrestrial areas (e.g. hedges, natural areas)		
	Seed/propagule dispersal	+	++	++	Yes	Insects, birds, mammals, fish and water
	Pest and disease regulation	+++	+++	+++	Yes	Non-target arthropods (beneficials, natural enemies), invertebrate and vertebrate predators and fungal species
	Climate regulation	++	+	+++	No	Several species (wild and domestic)
	Air quality regulation	++	+	+++	No	Plants
	Water regulation (quantitative aspects)	++	++ (acting as buffer zones)	+++ (acting as buffer zones)	Yes	Plants, microorganisms, soil fauna and beavers (dams)
	Erosion regulation		++	+++	Yes	Rooted plants soil fauna (ecosystem engineers)
	Natural hazard regulation (other than water regulation, e.g. avalanches and landslides)	+	+	+++	No	Rooted plants (shrubs and trees)
	Invasion resistance		+	++	Yes	Autochthonous species with a similar niche than invasive species
	Water purification/soil remediation/waste treatment	+	++	++	Yes	Plants, fauna, macrofauna, bacteria and fungi
Cultural services	Spiritual and religious values	++	++	++	Yes	All species
	Education and inspiration	+++	+++	+++	Yes	All species
	Recreation and ecotourism	++	++	+++	Yes	Fish (sport fishing), attractive plants and

Ecosystem service		In crop	Off crop		Strongly impacted by pesticides (direct or indirect effects via trophic interactions)	Potentially impacted taxa
			Small edge of field margins	More remote terrestrial areas (e.g. hedges, natural areas)		
						vegetation, vertebrates (bird watching, hunting) and attractive invertebrates
	Cultural heritage	+ to +++ (in traditional landscapes)	+ to +++ (in traditional landscapes)	+++	Yes	Preservation of structures constructed and/or modified by man and their typical biota
	Aesthetic values	++	++	+++	Yes	All species, in particular plants, vertebrates, attractive invertebrates and red list species
	Sense of place	++	++	++	No	Trees, patches of vegetation and ecosystems as landscape features
Supporting services (to produce other ecosystem services)	Primary production	+++	+++	+++	Yes	Algae and vascular plants
	Photosynthesis	+++	+++	+++	Yes	Algae and vascular plants
	Provision of habitat	++	+++	+++	Yes	Ecosystem engineers (e.g. beavers, earthworms, plants) and larger plants and animals that provide surfaces for periphytic organisms (e.g. shells of mussels)
	Soil formation and retention	++	+++	+++	Yes	Soil fauna (mainly ecosystem engineers e.g. earthworms, ants), plants (e.g. organic matter and peat

Ecosystem service	In crop	Off crop		Strongly impacted by pesticides (direct or indirect effects via trophic interactions)	Potentially impacted taxa
		Small edge of field margins	More remote terrestrial areas (e.g. hedges, natural areas)		
					formation)
Nutrient cycling	++	+++	+++	Yes	Microorganisms, primary producers, grazers, detritivores, consumers, predators
Water cycling	++	+++	+++	Yes	Plants and terrestrial and aquatic ecosystems

The following key ecosystem services are considered relevant in agricultural areas and are driven—amongst others—by NTTPs.

2.2.1. Non-target terrestrial plants as drivers for supporting services linked to nutrient cycling and water regulation

In the Scientific Opinion (EFSA PPR Panel, 2010), the in-field as well as the off-field areas are recognised as important spatial boundaries where supporting and regulating services take place.

In this opinion, it is acknowledged that nutrient cycling and water regulation are driven by non-target plants, in particular in the off-field areas.

2.2.2. Non-target terrestrial plants as drivers supporting food webs (primary production, provision of habitat and food for other non-target organisms, e.g. arthropods, birds)

The central position of non-target plants as primary producers supporting organisms at higher trophic levels has been repeatedly recognised in the Scientific Opinion of EFSA (EFSA PPR Panel, 2010). These supporting services include "primary production", "habitat provision" and "food for other trophic levels". "Primary production" and "habitat provision" are being driven by vascular plants in both in- and off-field areas, whereas the provision of "food" is related to off-field non-target plants (e.g. wild fruits) and in the in-field areas mostly to crop species. However, in-field non-target plants can be a food source for herbivorous birds and mammals but they have a predominant key role as a food source for non-target arthropods, which are themselves an important food resource for farmland birds and small mammals. Therefore, we integrate in-field non-target plants as a key driver of "food" into the set of SPGs and this also has to be addressed in the risk assessment scheme.

Several reviews and comprehensive studies highlight the key function of in-field plants and arthropods as food sources of farmland bird species, demonstrating long-term impacts of PPP application on populations of farmland birds (Campbell and Cooke, 1997; DEFRA, 2005; Bright et al., 2008, Jahn et al., 2014). Regarding the most intensively studied species, the grey partridge, a relationship between pesticides, food availability, breeding performance and population size has been demonstrated, with herbicides the main determining pesticide group (Marshall et al., 2001). It has been demonstrated that these unwanted effects on partridges were due to indirect effects of the PPP (i.e. intended primarily to control pest insects and weeds). They are nevertheless relevant for the risk assessment and management under Regulation (EC) 1107/2009, as according to the new data requirements (annexes to SANCO/11802/2010 Rev. 7) "the potential impact […] on biodiversity and the ecosystem, including potential indirect effects via alteration of the food web, shall be considered".

2.2.3. Non-target plants as drivers for the provision of genetic resources as well as for educational, recreational, aesthetic and intrinsic values

The ecosystem services listed here are provisioning and cultural services driven by non-target plants in agricultural landscapes. According to the EFSA PPR Panel opinion (2010), the aesthetic/recreational value of off- and in-field NTTPs (arable plant species) is a relevant ecosystem service. This holds the same for the genetic resources of in- and off-field NTTPs. Some of those species are considerably endangered and consequently need special attention.

2.3. Specific protection goals driven by ecosystem services provided by non-target terrestrial plants

It is proposed that SPGs be defined for non-target plants as drivers for:

- supporting (nutrient cycling) and regulatory (water regulation) services at the spatial scale edge of field up to the landscape (off-field area)—considering also protected areas;
- provisioning (food web and biodiversity/genetic resources) and cultural services for the spatial scale field (in-field area) and edge of field up to landscape scale (off-field area).

The magnitude of effects that are considered acceptable has to be differentiated and coupled to the area of assessment.

2.3.1. Specific protection goals for off-field NTTPs as key drivers for nutrient cycling, water regulation, food web support, aesthetic values and genetic resources (biodiversity)

For NTTPs in the off-field area, it is possible to define an SPG that integrates structural (genetic diversity, local abundance of species) as well as functional aspects of biodiversity (primary production, nutrient cycling, water regulation, provision of habitat and food). Owing to ecological redundancy, structural endpoints are generally more sensitive to PPP application than functional endpoints. Thus, effects at the population level of NTTP species should drive the risk assessment to make sure that a suitable level of protection for off-field NTTPs is ensured.

The protection goal for higher terrestrial plants aims to protect both plant species abundance (e.g. numbers and/or cover of individuals for single species) and plant diversity in an agricultural area. It is assumed that the biodiversity is maintained when the plant populations will not be affected, even for a short period, by the use of PPPs.

The SPGs for off-field NTTPs are defined as follows:

Ecological entity:	population
Attribute:	survival/growth/reproduction, abundance/biomass
Magnitude:	negligible
Temporal:	not applicable
Spatial scale:	edge of field
Degree of certainty:	high

2.3.2. Specific protection goals for in-field NTTPs as key drivers for food web support (primary production, provision of habitat and food for other non-target organisms, e.g. arthropods, birds)

Since the function of non-target plants as a food source is more relevant in this context than structural endpoints (plant diversity), the SPG should be aimed at the conservation or restoration of those functions as food or habitat sources rather than at the protection of the populations of single species. The functional group for food web support provides food (biomass of green material and seeds) and habitat (cover, host plant) provisioning for higher trophic levels. It was not possible to define the SPGs quantitatively but in sections 2.4 and 2.5 options are described. The SPGs for in-field NTTPs as key drivers for food web support are defined as follows:

Ecological entity:	functional group food web support (e.g. leafy crops, grass, seeds)
Attribute:	biomass for food web support
Magnitude:	negligible (landscape) to medium effects (field)
Temporal/Spatial scale:	weeks (no to few days during breeding/chick phase)
Spatial scale:	field/landscape
Degree of certainty:	high

2.3.3. Specific protection goals for in-field NTTPs as key drivers for aesthetic values and genetic resources

According to the EFSA PPR Panel opinion (2010), in-field NTTPs (arable plant species) are of a high aesthetic and recreational value and provide relevant genetic resources. The SPGs for in-field NTTP as key drivers for aesthetic values/biodiversity are defined as follows:

Ecological entity:	population/meta population
Attribute:	survival/growth/reproduction, abundance/biomass
Magnitude:	medium (meta-population), large effects (population) (both in-field), negligible (landscape)
Temporal:	not applicable/day to weeks
Spatial scale:	field/landscape
Degree of certainty:	high

2.3.4. Endangered species

In situations where endangered species are living in certain areas (including in fields) special measures have to be taken.

Ecological entity:	individuals/population
Attribute:	survival/growth/reproduction, abundance/biomass
Magnitude:	no effects
Temporal:	not applicable
Spatial scale:	field
Degree of certainty:	high

2.4. Mapping specific protection goals to test endpoints

In Table 2, a number of test endpoints are proposed for the risk assessment of NTTPs. The endpoints listed refer to situations where they can be derived directly from appropriate testing:

- Reproductive endpoint: based on the 5th percentile of distribution of effect rate $(ER)_{repro10}$ values;
- Biomass endpoints: one based on ER_{veg50} values, one based on ER_{veg10} values. Both the 5th percentiles of their respective distributions;
- Visual endpoint (e.g. chlorosis or bleaching): based on a 5th percentile of the distribution of $ER_{visual50}$ values.

The 5th percentiles should be calculated with the method that takes sample size into account, for example the methog provided by van Vlaardingen et al. (2004).

Since the tests for reproduction are not defined as a legal data requirement, a further option is provided to extrapolate reproductive endpoints based on the provided data according to the legal data requirements for NTTPs. If some of the data (e.g. $ER_{repro}x$) are not available to calculate the 5th percentiles, extrapolation methods can be used (see Appendix A).

For example, when no $ER_{repro10}$ values are available, the 5th percentile of the ER_{veg10} (when available) or the ER_{veg50} should be used in combination with a suitable extrapolation factor (EF). The EFs used to extrapolate from vegetative to reproductive endpoints and calculated with a 95 % certainty are 3(ER_{veg10}) and 35(ER_{veg50}), respectively. It should be noted that these EFs are not additive, i.e. the extrapolation from ER_{veg50} to $ER_{repro10}$ is direct and the EF of 35 incorporates both the difference in ER_x and the extrapolation from vegetative to reproductive endpoints.

Alternatively, when no ER_{veg10} values are available to calculate the 5th percentile of the distribution, the 5th percentile of the ER_{veg50} values can be used in combination with an EF (for 95 % certainty this

would be 34). Please note, this value cannot be used in combination with the abovementioned EF value of 3.

Moreover, these proposed EFs only cover the uncertainties with respect to the extrapolations from vegetative to reproductive endpoints, and they do not directly relate to the level of protection provided in the final risk assessment. The final risk assessment must also take into account an assessment factor applied on an endpoint to consider all remaining uncertainties (e.g. for a hazardous concentration $(HC)_5$ ER_{10}, remaining uncertainties include extrapolations from laboratory to field in terms of environmental stressors, of single-species tests to multi-species, etc.).

Where ER_{10} values are proposed for use in risk assessment, they are considered as a better representation for negligible effects than no observed effect rate (NOER) values.

Table 2: Overview of specific protection goals

Ecosystem service	Type of endpoints used in risk assessment[10]		Remarks	Consequences that may occur if specific protection goal is not achieved
(A) Specific protection goals for off field				
Biodiversity	Reproductive (long term)	Reproductive endpoint:[11] HC_5 $ER_{repro10}$	Negligible effects on reproduction at the edge of the field	Decline of biodiversity
	Vegetative (short term)	Biomass: HC_5 for ER_{veg10}	Negligible to small effects on biomass at the edge of the field Maintenance of plant species diversity may be hampered by direct impairment of reproduction (sexual and vegetative) as well as by indirect effects owing to competitive interactions in the field resulting from effects on growth, which is not covered by the reproductive endpoint	Decline of biodiversity
Nutrient cycling	Biomass and/or reproductive (long term)	Biomass (HC_5 for ER_{veg10}) and/or reproductive endpoint (HC_5 $ER_{repro10}$)	Some species are very important for nutrient cycling, e.g. legumes Mycorrhizas are important and therefore a wide variety of plants should be available Different plants have different chemical composition of their leaves and stem, etc., and decompose at different rates, which influences nutrient cycling Remark: it is generally not known which species are the key drivers for nutrient cycling, but the most abundant species are likely to be critical. Therefore, as a starting point, a conservative approach (HC_5 of a species sensitivity distribution calculated on a sensitive endpoint, e.g. $ER_{repro10}$) could be used. The risk assessment could be refined if more information becomes available. The risk assessment could then be focused on the relevant species and biomass impacts on these species	Decline of biomass of key populations will affect the potential for nutrient cycling

[10] Proposed assessment endpoints have to be combined with an appropriate assessment factor to cover all uncertainties linked to a risk assessment based on lab testing; these include, for example, lab to field, multiple exposure, crop to non-crop and so on (see section 2.1.3 for details).

[11] Reproduction tests are not foreseen as data requirements thus far and cannot therefore be required by the authorities. Therefore, options have to be provided to extrapolate the reproductive endpoint on the base of the tests according to the data reqirements (ER_{veg50}).

Ecosystem service	Type of endpoints used in risk assessment[10]		Remarks	Consequences that may occur if specific protection goal is not achieved
Water regulation	Vegetative (short term)	Biomass: HC_5 for ER_{veg50}	Small to medium effects could probably be tolerated before affecting water regulation Large changes in plant cover are likely to influence water regulation and in some cases there could be severe long-term effects if the topsoil layer is washed away. This will also adversely affect water courses	Decline of biomass will affect the potential for water regulation
Food web support	Vegetative (short term)	Biomass: HC_5 for ER_{veg10}	No effects at the edge of field No data are available to suggest magnitude of effects and to make a quantitative link to effects on food web	It is likely that the food web will be affected if there are effects on plant biomass
	Reproductive endpoint (long term)	Reproductive endpoint:[10] HC_5 $ER_{repro10}$	No effects at the edge of field Probably not applicable to all plants but important for specific plant species. Species on which other animals depend on for food or reproduction	It is likely that other organisms, depending on the availability of nectar and pollen, and particular species as host plants will be affected when these are unable to flower or maintain a viable population
Aesthetic values	Visual effects (e.g. chlorosis)	Visual endpoint: HC_5 $ER_{visual50}$	Slight and temporary chlorosis (bleaching) may be considered acceptable as long as they do not last longer than a few days	Parts of the agricultural landscape will be unattractive for a short period of time
	Reproductive endpoint (long term)	Reproductive endpoint from appropriate tests:[10] HC_5 $ER_{repro10}$	Small and temporary effects may be considered acceptable. Probably not applicable to all plants but important for specific plant species	Highly valued flowering plants from an aesthetic point of view will be less visual or will be disappearing from the agricultural landscape
Genetic resources	Reproductive endpoint (long term)	Reproductive endpoint:[10] HC_5 $ER_{repro10}$	Small and temporary effects at the edge of field could be tolerated but there is not enough information available to make a quantitative link between effects at the edge of field and landscape level	Decline of genetic variability and therefore the capability of coping with stress caused by the use of plant protection products or other stressors
(B) Specific protection goals for in field (see SPG chapter SPG B–D)				
Food web support	Vegetative (short-term and/or reproductive endpoints)	Biomass and/or reproductive endpoints	The use of herbicides and/or compounds with herbicidal activity will influence the food web in fields. Information on how much of the food web should be maintained has to come from other risk assessment schemes (e.g. arthropods, birds, etc.) Appropriate solutions could be set aside areas, untreated buffer strips or other types of in-field strips	Plants in field may be important to support the food web. Removal of non-target (in the case of non-herbicide applications) and also target plants (in the case of herbicides) will most likely impact food web support

Ecosystem service	Type of endpoints used in risk assessment[10]		Remarks	Consequences that may occur if specific protection goal is not achieved
Aesthetic values	Visual endpoints (e.g. chlorosis or bleaching) or reproductive endpoint (long term)	Visual endpoints (e.g. chlorosis or bleaching) HC_5 for visual endpoint $ER_{visual50}$, reproductive endpoint[10] HC_5 for $ER_{repro10}$	Slight and temporary chlorosis may be considered acceptable as long as they do not last longer than a few days Small and temporary effects on reproduction endpoints may be considered acceptable	Highly valued plants for an aesthetic point of view will be less visual or will be disappearing from the agricultural landscape (i.e. inside the cropped area) and parts of the agricultural landscape will be unattractive for a short period of time
Genetic resources	Reproductive endpoint (long term)	Reproductive endpoint:[10] HC_5 $ER_{repro10}$	Small and temporary effects could be tolerated. This part of the table is particular important for rare species that are mainly living in the treated agricultural areas	A decline of genetic resources will possibly occur and will probably influence the ability to cope with other stressors
Endangered species	Reproductive endpoint (long term) and in particular cases it could be necessary to protect individuals of the species	Reproductive endpoint and when necessary any endpoint that will be needed to protect the species as such	In the case of endangered species living in certain areas (including arable land), special measurements should to be taken More guidance will be provided by a working group of the Scientific Committee of the EFSA in 2015	Decline of the endangered species and ultimately the extinction of the species

In Table 2 it is proposed in some cases to allow small to medium effects for a short period but there is currently no guidance for the implementation of such measures. This will partly depend on the recovery capabilities of the ecosystem considering the spatiotemporal environmental conditions such as other stressors or pesticide application schemes and landscape characteristics. The environmental working group of the Scientific Committee of EFSA will at the end of 2015 provide guidance on how to deal with recovery (EFSA mandate number M-2013-0098). Furthermore, the recovery option also requires higher tier approaches.

Whether endangered species are covered by the criteria used or whether a compound is environmentally safe under a particular restriction is at this moment assessed by the environmental working group of the Scientific Committee of EFSA (EFSA mandate number M-2013-0098).

2.5. Implementation of the specific protection goals in risk assessment using operational protection goals

The general protection goal for higher terrestrial plants is to maintain the biodiversity of plant species in an agricultural area. The risk assessment procedure should ensure that the use of PPPs does not have any long-term repercussions for the abundance and diversity of non-target plant species. It is assumed that this is achieved when:

(a) the community of non-target plant populations in the off-field area is not affected by the use of PPPs, and (section 2.5.1)

(b) the populations of non-target plant species in the in-field area are adequately managed to achieve the defined SPGs. (section 2.5.2)

2.5.1. Realising the specific protection goal for the off-field situation

With respect to (a), above, the edge of the field, i.e. where the off-field starts, will be considered as the relevant spatial boundary (see also section 2.7) to achieve a sufficiently protective risk assessment, although larger spatial scales in an agricultural area may also be considered appropriate to protect the community of plant species in the off-field area. Currently, no appropriate scientific basis or methods are available for proposing a landscape level approach that would be sufficiently protective.

There is some information that supports the hypothesis that laboratory studies are more conservative than real exposure in the field (see section 6.1.3) if the plants are grown individually in pots. However, other studies show that sensitivity could be higher in multi-species situations (i.e. typically under field situations, owing to, for example, interactions between plants or between different plant species); in this context laboratory studies may not be more conservative (see section 6.1.3). In relation to the SPGs for "off-field", a thorough risk assessment scheme can be provided on the basis of current data requirements. It is assumed that the SPG can be achieved if it is based on current data requirements and a realistic worst case exposure level (e.g. the 90th percentile of downwind drift values) and the effect value (e.g. HC_5 of ER_{10}) in combination with an appropriate assessment factor will exclude unacceptable risks for off-field plant populations.

The standard endpoint from the higher plant toxicity tests is an ER_{50} value for vegetative endpoints, which in most cases is based on biomass. However, as the intention is to protect plant populations, it is advisable to also include reproductive endpoints. It is also advisable not to use an ER_{50} value because 50 % of the biomass of tested species would then be affected by the PPP. In order to ensure that there are no effects a NOER would therefore be preferable and an ER_{10} value can be used as a surrogate NOER value.

It is therefore proposed to use the 5th percentile of the ER_{10} toxicity data (species sensitivity distribution) and the 90th percentile of the downwind exposure distribution (traditional exposure level according to FOCUS (2001)). Based on the assumptions made above, the resulting operational protection goal (OPG) can be described as follows: 95 % of the plant species will not be exposed

above their ER_{10} under consideration of realistic worst case off-field scenarios (e.g. 90 % of the calculated exposure distribution in the defined scenario).

It is important to note that this OPG is based on a limited number of single-species tested in the laboratory, whereas the protection goal "negligible effects on plant populations at the edge of the field" refer to the field situation. Therefore, the remaining uncertainties, need to be addressed by higher tier testing or by an assessment factor, which will cover for extrapolations from single species to multiple species, from laboratory to field, and the protection of the remaining 5 % of the species, etc. Establishing the size of the assessment factor is a risk assessment task and needs calibration with the reference tier for a sufficient number of plants and herbicidal compounds.

There are two aspects of the use of PPPs that are not covered under the legislation at the moment, i.e. the use of compounds in tank mixtures (including the influence of additives) and the sequential use of different PPPs during a growing season (see sections 8.1 and 8.2). Therefore, it is not possible to guarantee that the protection for the off-field situation will be achieved unless the actual use of PPPs in crop protection programmes is simulated in ecotoxicity tests.

2.5.2. Realising the specific protection goal for the in-field situation

The in-field area covers the cropped and non-cropped area owned by the farmer that might be treated with a PPP (see section 2.7). SPGs for non-target plant species are also described in section 2.1 for the in-field area, especially regarding (1) the function of NTTPs as drivers for the provisioning service "food web support" for other non-target species at higher trophic levels, (2) their intrinsic ethical or aesthetic and recreational values and (3) endangered species.

When assessing the risk of non-target plants exposed to PPPs considering the SPG "food web support" for non-target arthropod and farmland bird species, for example, it has to be considered that the relative importance of plant populations in arable land as SPG drivers will strongly depend on the availability of alternative food sources in the landscape. Thus, the severity of the effects of PPP use on the conservation status of populations of farmland bird species and of other non-target organisms will strongly vary, being more pronounced in intensive agricultural areas with a scarcity of suitable refugial habitats than in highly diverse landscapes (e.g. structure and diversity). Given the diversity of production areas in Europe's regulatory zones, the definition of one representative landscape best describing the risks arising from PPPs use on the SPG "food provision" by non-target plants is unrealistic.

Regarding the SPGs "aesthetic and recreational value" and "endangered species", Member States do recognise the need for protecting non-target plant species in in-field areas. Management schemes apply at national or regional scales, including measures such as economic incentives for no-spray zones or flowering margins, partly as rural development measures in the frame of the Common Agricultural Policy. However, the current EU agricultural and environmental policy lacks concrete requirements regarding the protection of non-target plants in the field. For the time being, the definition of the target level of protection of non-target arable plant species in the context of Regulation (EC) 1107/2009 is up to individual Member States.

However, the guidance document should enable risk assessors to determine whether or not the application of a certain product is deemed harmful to the populations of non-target plants in field, and whether this is deemed to harm higher trophic levels. Possible risks should be characterised and management options proposed to reduce identified risks and thereby enable risk managers to decide on appropriate risk management measures, depending on the intended level of protection for non-target arable plants—considering specific ethical, social or economic aspects—as well as on the intended level of protection for plants, with regard to the SPG "food web support".

Herbicides, plant growth regulators and compounds with herbicidal activity will affect the plants in the agricultural area and will therefore affect the food web and shelter for a large number of species (e.g.

birds, mammals and insects) and it will also affect the aesthetic and recreational value as well as the genetic resources of the landscape.

Available information from biodiversity monitoring activities (e.g. on population trends for farmland birds or diversity of rare arable weeds) could be considered as important indicators for the long-term impact of PPP use on the SPGs "ethic and aesthetic value", "endangered species" and "food web support".

The results of the overall assessment should enable risk managers at the individual Member State level to decide on the acceptability of the described effects on the SPGs "food web support", "aesthetic values", "genetic resources" and "endangered species" considering the specific ecological and agricultural situation in the authorisation zone. Risk mitigation measures for the management of the described effects are available and include, for example, compensation measures, especially cropped no-spray zones, fallow land and flowering margins. For a detailed description, refer to section 2.13.

2.6. Species and growth stages requiring particular consideration

Species and growth stages requiring particular consideration can be the most sensitive species to an array of herbicides; they may be rare species threatened by herbicide misplacement, they may be underrepresented species in herbicide testing schemes (e.g. ferns, mosses, liverworts, hornworts, horsetails, lichens and woody species) or they may be sensitive at phenological stages untested in current regulatory guidelines (reproduction, whole life cycle). Studies indicate that no one species is sensitive to all herbicides and both crop and non-crop species seem to be equally sensitive (but not always; see section 3.1) depending on herbicides used and life stages investigated. It has been shown that reproductive endpoints are more sensitive than vegetative endpoints, and that they are easier to measure in non-crop species than in crops which are often not grown for their seed production, e.g. onions, potatoes, cabbage.

In general, only crop species are tested for herbicidal effects and they are assumed to represent all non-target plant species. Crops are usually annual species belonging to a restricted number of families (see Schmitz et al., 2013b). Although the limited data available seem to suggest no difference in sensitivity between annuals and perennials (in short-term vegetative tests), it may be that annual species with short-lived seed banks are more at risk because they need to reproduce yearly (most rare arable weeds are annual species). Very limited data are available on herbicide sensitivity of ferns, mosses, liverworts, hornworts, horsetails and woody species, which may be considered at risk (see section 3 on the selection of species). Likewise, generally, data are scarce on the susceptibility of various endpoints, including effects on flowering, pollen viability, seed production and more generally on the whole life cycle.

Species and growth stages requiring particular consideration may also refer to key plant species for particular organisms (e.g. *Erodium cicutarium* and *Geranium* spp. are the favoured foodplant of the brown argus (*Aricia agestis*); the Queen of Spain Fritillary (*Issoria lathonia*) larvae feed only on *Viola arvensis* and *V. tricolor* and adults mainly feed on flowers of *Jasione montana*; the small copper (*Lycaena phlaeas*) larvae feed only on *Rumex* spp.; and, in North America, *Asclepias* spp. are essential for the monarch butterfly). Species and growth stages requiring particular consideration can also be the plant species dominating or at the base of a given food web (e.g. *Typha* spp. dominating wetlands).

Many arable weeds have become rare to the extent that they are registered in the Red Data Books (International Union for Conservation of Nature) of several countries. These species can be considered vulnerable since they are also closely associated with cropland areas, and thus are subjected to recurrent herbicide spraying.

2.7. Spatial boundaries

Specific protection goals need to be defined for in- and off-field areas (section 2.1). The buffer strip is located in field and has the same protection goals as the in-crop area as well as the functions to

mitigate exposure of the off-field area (drift and run-off reduction) and may serve as a reservoir for recolonisation of the in-field area if there is no suitable off-field habitat. The quality and size of off-field habitats is important for the maintenance of non-target organisms in the agricultural landscape. Populations of non-target organisms in areas with sufficient off-field habitats may tolerate a greater impact from pesticide use than populations in areas with little off-field habitats. In order to account for this in the risk assessment, it would be necessary to have detailed information on the actual off-field habitats in the different landscapes and crops. Such information is currently not available. To overcome this problem and to aid the development of a generic risk assessment scheme, it is proposed that generic off-field protection goals are defined which are independent from the actual type of off-field habitat of individual fields.

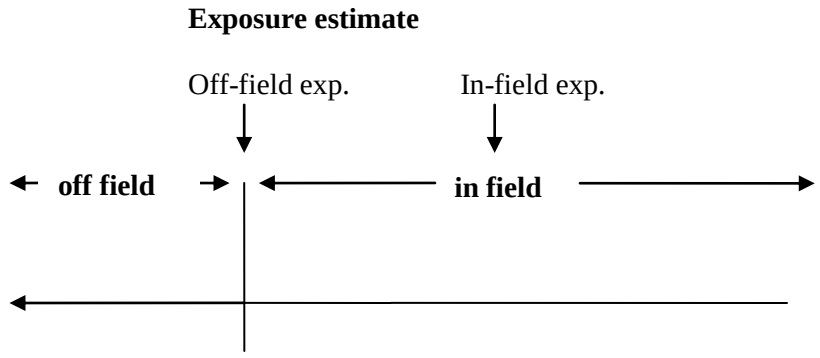
The size of buffer strips which are needed in order to mitigate the risk to non-target organisms depends on the agricultural landscapes context. Larger buffer zones may be needed in landscapes with little (semi-)natural off-field habitats. If the consequence of implementing buffer zones is the creation of larger fields with less off-field habitats then this could have an impact on biodiversity. The working group cannot address this problem but risk managers should be made aware that other measures are needed to maintain biodiversity in agricultural landscapes which are outside the remit of the pesticide risk assessment. The implementation of buffer zones is only one risk mitigation option and others could be developed.

It is necessary to define the spatial boundaries of the off-field area and the way the emission is translated to an exposure in the off-field area. These boundaries relate to the protection goal in relation to the route and distance covered of the emission coming from the in-field area. The choice of such a distance will be the result of both scientific (e.g. is there a critical maximum area that can be at risk without affecting the population of interest?) and regulatory decisions (is that distance acceptable from a regulatory point of view?).

Predicted environmental concentrations (PECs) could be provided for different distances from the field boundary and choices need to be made depending on the crop, group of non-target organisms and their SPG. This PEC calculation allows definition of buffer strips and the risk assessment in the off-crop area at the same time.

2.7.1. First step

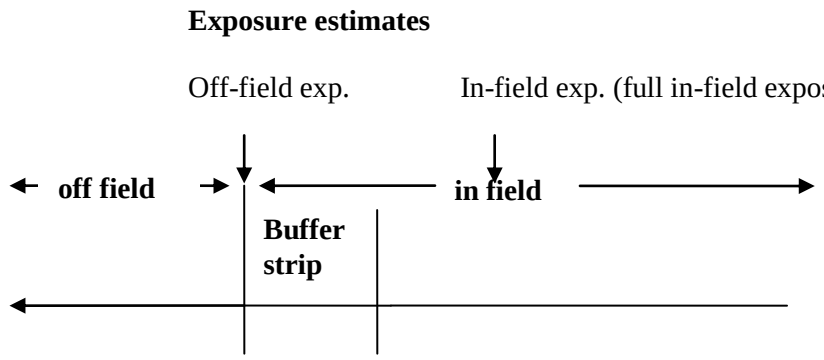
The exposure and risk is assessed for the in- and off-field areas. If the SPGs of in- and off-field areas are met, no further risk assessment or risk mitigation is needed.



2.7.2. Second step

A buffer strip is necessary if the off-field protection goal is not met in the previous risk assessment. The buffer strip is in field. The maximum tolerable exposure to meet the off-field protection goal is calculated. The width of the buffer strip is calculated on the basis of this maximum tolerable exposure estimate, the required reduction factor and the reduction potential of the buffer strip. For example, for

spray drift, the final width of the buffer strip depends on the combination of the height of the vegetation in the buffer strip and drift reduction techniques. If, for example, a wind break is in the in-field area, then the drift to the off-field area is significantly reduced compared with a buffer strip without vegetation. A table on the reduction of spray drift from the combination of spray drift nozzles and the width of the buffer zone can be found in Huijsmans and van de Zande (2011).



The various elements sketched in the scheme are defined as follows:

In-field: piece of land for cultivation with crops, managed typically by one farmer, with a distinct boundary

Buffer strip: in-field; non-treated cropped or non-cropped zone of a defined width at the edge of a field that is influenced by the farmers action (e.g. spray drift). The buffer strip is normally enforced by authorities and underlies prescribed actions in order to meet the off-field SPG. In addition, buffer strips may provide a recovery potential for the cropped area

Off-field: area surrounding a field; (semi-)natural habitats with high ecological value such as hedgerow, grass strip or simple structures (fence or a bare strip of land); normally no short-term changes in vegetation, in most cases not to be influenced by the farmer. Off-field also includes man-made structures, e.g. an adjacent field, roads, etc.

2.7.3. Proposed spatial scale for risk assessment

The initial assessment should calculate the acceptable concentrations in the off-field area. From this, the distance from the last row of treated crop at which the off-field protection goal is met can be back-calculated.

The risk assessment does not assume a pre-defined distance to the off-field area. The exposure assessment starts at the field edge and calculates the distance at which the off-field protection goal is met. Let us assume that the regulatory acceptable concentration is equal to the amount of active substance at a distance of 5 m. The full risk mitigation equivalent to a 5 m buffer zone needs to be achieved in the in-field area. Standard options for reducing the width of the in-field buffer strip could be provided in the risk assessment, e.g. vegetation in the buffer strip of a certain height or wind breaks or drift reduction nozzles. The risk manager decides whether the risk is manageable under the national conditions to achieve the required reduction of exposure in the off-field area (e.g. considering agricultural practice or national policy on implementation of buffer zones).

With this approach, it does not need to be assumed that the off-field protection goal needs to be met 1 m and 3 m away from the last row of the treated crop, as was previous practice. Thus, a risk management decision can be taken based on knowledge of how much distance is kept to the edge of the field (it may be different in Member States and crops) and based on national policies for implementation of buffer strips, e.g. obligatory vegetated buffer strips of a certain width.

2.8. Protection of in-crop/off-crop plants from pesticide application—rare arable weeds

The PPR Panel acknowledges that the risk assessment for NTTPs growing within crops cannot afford the same level of protection from pesticide application than off-crop plants. This is because the aim of herbicide use in agriculture is to control weeds in order to optimise crop productivity. Thus, non-crop species growing within cropland include both weed species that interfere with crop yield and some rare species that may be of conservation value (see also section 2.1). Risk assessment should consider species for in- and off-field areas. However, the level of risk that is acceptable will generally be higher within crop fields owing to the need to control weeds.

It is known that some rare arable weeds preferably, or only, grow in close association with crops; therefore, the question arises about how rare arable weeds can be protected. This is a difficult issue since those rare arable weeds were problematic species to be eradicated by farmers, and recent management practices have been very successful. Arable plants are generally annual ruderal species that grow where there is regular disturbance (although arable plants also include some problematic perennial species), and in the case of arable weeds, the disturbance is provided within agricultural fields. There is a good indication that a large number of annual arable weed species are under threat of extinction in several European countries (Denmark, Germany, Spain, France, the Netherlands, Sweden, the UK and Turkey; see below). It is recommended that measures are taken to compensate effects of herbicides on rare arable weeds. Some suggestions are included in section 2.12.

Many arable weeds have become rare owing to intensive management practices introduced in the last 50 years: extensive use of agrochemicals applied with ever increasing machinery size, increased field size and destruction of marginal habitats for the use of this machinery, better seed cleaning, use of high-density crop shading out weeds, and other modifications in crop types and management such as monoculture and timing of harvest (Sutcliffe and Kay, 2000).

2.8.1. Historical data

A few plant surveys are available comparing historical and current data, which demonstrated a shift in species composition in agricultural fields in the last 50–60 years. Potts et al. (2010) used in-crop field survey data collected annually since 1968 in 106 fields in the Sussex downs in the UK. Data from previous surveys in the same locations were also used to assess species assemblages prior to herbicide use (1946) and from the start of herbicide use (1946–1968). The objective was to record weed species occurrence and assess general weed abundance. Fifteen species and one subspecies were lost from the flora (out of 217) between 1946 and 2005. There were 15 gains during the same period. There was a higher occurrence of weed flora in fields where no pesticides were used. This study does not mention the status of the lost species, i.e. whether or not lost species were rare arable weeds. However, from the appendix provided by Walker et al. (2007), rare species could be identified, although the species list provided by Walker is probably not exhaustive. None of the species gained was classified as rare, while 12 of the species lost had been classified as rare by Walker et al. (2007). Potts et al. (2010) concluded that limiting the use of herbicides would help restore the flora in cereal ecosystems, providing that species are still available in the seed bank.

Another study also conducted in the UK noted the occurrence of rare arable weed species in a 30-year interval (Sutcliffe and Kay, 2000). Many of the species surveyed in the early 1960s were still found in 1997. However, their abundance had declined markedly. Several species such as thorow-wax (*Bupleurum rotundifolium*) and corncockle (*Agrostemma githago*) were not found and are probably extinct in Britain.

In the Thiérache region of northern France, a study was carried out with the objective of comparing a 19th century floristic list with cumulative records of surveys conducted in the second half of the 20th century (available database), including a thorough survey conducted between 1995 and 2000 (Van Calster et al., 2008). The landscape of the surveyed region had not been greatly modified during this period: forest, grassland and arable land cover remained the same and human population density slightly decreased. Nevertheless, the decrease in arable species richness was very pronounced (53 %)

compared with forest species (31 %). Most extinct species were already rare historically. The authors suggest that the use of large machinery, better seed cleaning techniques, a decline of crop types planted and the use of pesticides and fertilisers may be related to arable weed decline.

In the Côte d'Or region of Burgundy in the eastern part of central France, 158 fields surveyed between 1968 and 1976 were revisited in 2005–2006 (Fried et al., 2009). The aim of the study was to assess changes in weed communities and also to evaluate the importance of crop edges as potential refugia for threatened arable weeds. Results indicated that many species of conservation value had disappeared or seriously decreased but still persisted in crop edges.

Surveys of the weed flora conducted in Denmark in 1964–1970, 1987–1989 and 2001–2004 revealed an interesting pattern (Andreasen et al., 1996; Andreasen and Stryhn, 2008). There was a 60 % decline of 67 weed species between the first (1964–1970) and second surveys (1987–1989) (Andreasen et al., 1996), and only *Stellaria media* increased significantly and solely in grass ley fields. In sharp contrast, between the 1980s and 2000, weed frequency of the same 67 species increased on average by 45–75 % in annual crops and remained at the same level in grass ley (Andreasen and Stryhn, 2008). Nevertheless, many of the less common weed species in the 1987–1989 survey did not exhibit significant changes (*Anagallis arvensis* ssp. *arvensis* L., *Anchusa arvensis* (L.) Bieb., *Anthemis arvensis* L. and *Rumex acetosella* L.), and a few species became less common (*Elymus repens* (in winter wheat), *Plantago major* L. and *Scleranthus annuus* L. ssp. *annuus*). The authors attributed the rise in weed frequency to a reduced usage of herbicide owing to agricultural policies such as the withdrawal of several herbicides from the market, a 50 % reduction in spraying intensity per hectare, the conversion from conventional to organic farming and the declining use of commercial fertiliser (use of manure remains the same).

In the Netherlands, 53 of the 190 weed species were classified as rare (Kleijn and van der Voort, 1997; see also de Snoo, 1997). In Germany, approximately one-third of the 250–300 weed species were considered rare (Eggers, 1987; see also Albrecht and Mattheis, 1998). In Britain, 54 species in the arable flora were classified as rare in the early 1990s (Wilson, 1994). It increased to 97 in the following decade (Wilson and King, 2003). These species can no longer survive in fields because of high weed control efficiency and are thus relegated to the extreme edge of crop fields or in-field margins.

Modifications of the farmland flora have been documented in many other European countries. Several weeds were shown to have declined in Spain (Romero et al., 2008; José-Maria et al., 2010), Sweden (Fischer and Milberg, 1997) and Turkey (Tùre and Böcük, 2008).

2.8.2. Surveys showing preference habitats for rare arable weeds

Arable weeds, including rare arable weeds, are more common at the edge of fields and field corners where weed control is not practised or is less efficient than at the centre (Marshall, 1989a, b, 2009; Walker et al., 2007; Romero et al., 2008; José-María et al., 2010). Arable weeds need some form of regular soil disruption, e.g. tilling; therefore, hedgerows and boundaries with little disturbance or sown with grasses and which favour perennial species are not good habitats for rare arable weeds. Wilson and Aebischer (1995) in the UK conducted three surveys in cereal crops in 1988 and 1989 at 10 distances from the crop edge. Most species were found or were more abundant in the first 4 m from the edge within the crop. In Spain, it was found that organic practices favoured rare arable weeds and that weed diversity was concentrated in the crop edges (Romero et al., 2008). Although weed diversity was lower in conventional farming, especially in field centre, there was only a marginally significant difference in species richness between organic and conventional at the field edges. José-Maria et al. (2010) incorporated landscape variables in their analysis (percentage of arable land and human settlements) to farming type and intensity and found similar results. Plant species richness and composition were influenced by both scales of agricultural intensification, especially in field centres.

Several studies have evaluated the influence of different managements of field margins and boundaries on rare arable weed species. Marshall (2009) studied the influence of sown grass strips at the margin of fields on plant diversity, including rare arable weeds. It was found that sown grass strips increased plant diversity (mostly polycarpic species) and reduced unwanted weeds but did not enhance rare arable weeds and may even be harmful to them. The impact of crop management and agricultural land-use change on the European arable flora was assessed by botanical experts using the Red Lists of vascular plants in 29 European countries (Storkey et al., 2012). Wheat yield was used as a proxy for agriculture intensification. Countries with large national wheat yield (or substantial intensive agriculture) contained a higher proportion of rare or threatened arable plant species than countries with more modest wheat yield. The two most important variables explaining the decline in rare arable weeds were fertiliser and herbicide use. These findings point to agricultural intensification as the overall culprit for the decline of rare arable weeds (see also Bilz et al., 2011).

Initiatives to increase populations of weeds that are now considered as rare arable weeds have concentrated on field edges, including the establishment of conservation headlands, which consist, in the Netherlands, of a 2- to 3-m-wide linear strip at the edge of fields where no pesticides are used (Kleijn and van der Voort, 1997). In other countries, conservation headlands can be up to 6 m wide and compensation is provided to farmers for the loss of crop yield. The usefulness of this practice was evaluated in several countries and was deemed effective (http://www.arableplants.org.uk/assets/Agri-environment/ManSheet-Conservation-Headlands.pdf). Nevertheless, numerous studies have demonstrated that uncropped but tilled field edges were best for the conservation of rare arable plant species. Walker et al. (2007) evaluated the influence of uncropped cultivated (tilled) margins, spring fallow and cropped conservation headlands with or without fertiliser inputs and reduced pesticides on plant diversity, including 86 rare species. A conventionally managed cereal crop was used as a control site. The study was conducted in several areas of England. Species diversity and rare species richness were highest in uncropped cultivated (tilled) margins, followed by spring fallow and cropped conservation headlands with no fertiliser and reduced pesticide application. In cropped areas with no fertiliser and reduced pesticides, a fair number of rare arable species were found compared with control sites. A limited number of rare species were found in the control sites. However, the least frequent scheme followed by farmers was the uncropped cultivated margin. In a complementary study, Critchley et al. (2006) investigated the effect of various depths and timings of soil cultivation on species composition of annuals and dicotyledons in 6-m-wide uncropped field boundary strips. Generally, cultivation timing had a greater effect than depth. However, there was a high variability in response among the study sites owing to variation in plant species with different germination requirements.

All the studies mentioned above pointed to the fact that uncropped cultivated (tilled) margins appear to be best for rare arable weeds. Herbicides (and to a lesser extent, fertilisers) are very detrimental to rare arable weeds. However, uncropped cultivated margins are not a preferred option by the farming community and, consequently, are not practised much among farmers (Walker et al., 2007).

Although rare arable weeds need protection from pesticide use, this brief literature review revealed that their management should be considered in the light of the overall agricultural practices of crop margins (buffer zones or in-field non-treated strips, etc.).

2.9. Reproduction and seed bank: importance and herbicide effects

Herbicides will affect the species pool of a plant community by selectively decreasing seedling establishment or by reducing reproduction. The consequences of reducing plant species fitness is a concern for plant gene flow and the survival of a species. In many studies, it was shown that late season/autumn application of some herbicides (e.g. glyphosate, glufosinate ammonium, sulfonylureas, imidazolinones) reduced seed production and seed viability of plant species (Isaac et al., 1989; Clay and Griffin, 2000; Brewer and Oliver, 2007; Walker and Oliver, 2008; Gauvrit and Chauvel, 2010). A study conducted by Fletcher et al. (1993) in the USA revealed that the sulfonylurea herbicide chlorsulfuron caused a significant reduction in the reproduction of cherry plants at 0.2 % of the

recommended label rate, and this was more pronounced when the plants were at bud and flower stage development. One single application of picloram on a native shrub in the USA decreased flowering and seed set, and the effects lasted for four years (Crone et al., 2009). In Denmark, hawthorns in seven different hedgerows were sprayed with the sulfonylurea metsulfuron methyl at the reproductive stage (Kjær et al., 2006a, b). Effects were observed soon after the spray event and continued to be significant one year after the herbicide application.

A reduction of the plant reproduction has an effect on the seed bank. Soil seed banks are an important component of vegetation dynamics in ecosystems (Leck et al., 1989) and can be regarded as the "memory" of plant communities, especially those dominated by annual species (Cavers, 1995). Researchers have reported a steady decline in total seed bank densities in crop fields receiving repeated herbicides applications (Roberts and Neilson, 1981; Ball, 1992; Rokich et al., 2009). Conversely, omitting in-crop herbicide use in agricultural fields increased the seed bank density (Gulden et al., 2011. Likewise, the number of seeds in soil increases when converting from conventional to organic farming owing to mechanical weed control replacing herbicide applications (Albrecht, 2005). A high seed density was measured in the seed bank of organic fields (José-Maria and Sans, 2011).

Off-crop habitats adjacent to in-field areas harbour a high diversity of plants in the aboveground vegetation as well as in the seed bank (Boutin, 2006; Boutin et al., unpublished data). Studies have shown that herbicide drift can affect the aboveground vegetation present in field boundaries (Kleijn and Verbeek, 2000; Gove et al., 2007). It is likely that the seed bank composition and abundance in boundary habitats abutting crop fields are also reduced by recurrent herbicide drift, although to our knowledge there has been no study conducted on the impact on the seed bank in these habitats.

2.10. Plant recovery from sublethal doses of herbicide

The ability of young plants to recover when sprayed at sublethal doses of herbicides has rarely been studied. Reuter and Siemoneit-Gast (2007) showed that most plants growing in a microcosm community did not recover from a single herbicide application within 42 days. In a few cases, it was found that plants were sometimes able to compensate for initial losses of biomass, and some species were able to recover to levels comparable to the controls (Marrs et al., 1991a; Riemens et al., 2009). However, even if plants are able to compensate for biomass loss in the long term, measurements of biomass alone may not be enough to truly understand the effects on a plant's ability to reproduce in the future. Although recovery may occur in certain species, the potential energetic expenses required to recover may have negative effects on the reproductive success of the individual through reductions in seed outputs. This could be aggravated under natural conditions, where both intra- and inter-specific competition for resources as well as biotic (e.g. pathogens) and abiotic (e.g. nutrient availability) factors further reduce the ability to recover, especially in the presence of non-sensitive species. A plant that receives a sublethal dose may experience a reduction or delays in flowering or fruit production, which could potentially inhibit seed output that year if it fails to occur within the growing season. Such situations may lead to an oversimplification of the plant community or favour certain groups of species over other sensitive groups (e.g. grasses when broad-leaved herbicides are sprayed or perennial over annual species).

Carpenter and Boutin (2010) conducted a greenhouse study with 10 crop and 10 wild species. Half the plants of each species (eight doses + controls × six replicates) were harvested four weeks after being sprayed (short term). The remaining plants were harvested several weeks later, coinciding with seed set or natural senescence (long term). Total aboveground biomass and several endpoints related to vegetative growth and plant reproduction were measured. It was found that plants recovered in terms of total biomass over time, except for two species (*Phytolacca americana* and *Juncus dudleyi*). However, of the 12 species for which reproductive output could be assessed, in seven cases the reproductive endpoint measured was lower than the endpoints based on the short- or long-term biomass. Likewise, Riemens et al. (2008) observed larger effects of lower doses of glufosinate on seed production in *Stellaria media* than on aboveground biomass. A study presented in Schmitz et al. 2013b

revealed that the process of recovery after herbicide spray was species and herbicide dependent. The sensitivity of two of the six species tested, *Galium mollugo* and *Silene nutans*, increased over time, comparing effects at 14, 28 and 42 days. Effects on reproduction did not appear to have been investigated as no detail was provided regarding the phenological stage of the species when data were recorded.

Two additional studies were available to assess the effects and recovery of the reproductive output. The first study focused on the susceptibility of six test species, three annuals and three taxonomically closely related perennial species: *Geranium molle*, *G. robertianum*, *Silene noctiflora*, *S. vulgaris*, *Tripleurospermum inodorum* and *Achillea millefolium* to mecoprop, metsulfuron and glyphosate (Strandberg et al., 2012). In a second unpublished study (Strandberg, personal communication), an additional four species were assessed: *Festuca ovina*, *Agrostis tenuis*, *Solanum nigrum* and *Echinochloa crus-galli* were tested with glyphosate and foramsulfuron plus iodosulfuron. The herbicides were applied at two growth stages: the early growth stage was from three leaves to ten leaves and the late growth stage was at the flower bud stage. Plants from three replicates of each treatment (five to seven doses) were harvested three to four weeks later. Seed production was assessed at maturity with the remaining three replicates per dose. Results show that seed production was a more sensitive endpoint than vegetative measure, regardless of the time of spray, for all species where reproduction could be measured. These results (Carpenter and Boutin, 2010; Strandberg et al., 2012) indicate that the measurement of biomass 21 or 28 days after spray may underestimate toxicity for most species. They also indicate that an assessment of recovery over time is essential to fully assess the toxicity and impacts of a herbicide.

In the studies above, the time taken for plants to recover following herbicide impact was not measured. A further study was initiated with the objective of assessing the rate of recovery of a selection of wild species from two different communities, terrestrial and wetland habitats, following exposure to the herbicide chlorimuron ethyl (Carpenter et al., 2013). Nine terrestrial upland species (one monocot and eight dicots) and eight wetland species (three monocots and five dicots) were used in a greenhouse experiment. The experimental design included short- and long-term harvests as in previous studies. In addition to total aboveground biomass, several endpoints related to vegetative growth (i.e. height) and plant reproduction were measured during the course of the experiment in order to determine the time it took for plants to fully recover from sublethal herbicide effects. Results on total aboveground biomass revealed that plants in the long-term treatments had a higher effect concentration $(EC)_{50}$ (effective concentration that causes a 50 % decrease in biomass) values than those calculated in the four-week treatments, indicating a trend towards recovery. However, in three cases (all annual species), the reproductive output was a more sensitive endpoint than the short-term biomass. Of the 11 species that were affected and had a measurable reproductive parameter, three exhibited equal recovery of both the vegetative and reproductive parameters at equivalent doses by the end of the experiment, and one species displayed recovery of the reproductive stage at a lower dose than of the vegetative stage; the average recovery time was 7.5 weeks. In the case of the remaining seven species, recovery of the reproductive parameters never occurred (compared with control plants), while the vegetative measure showed more propensity to recover quickly from herbicide injury (Carpenter et al., 2013).

The delayed average flowering time of various plant species populations exposed to increasing doses of herbicides can have consequences. Flowering time (time to first flower) was impeded for most species as doses of glufosinate ammonium and chlorimuron ethyl (Boutin et al., 2014) increased. Likewise, flowering was reduced and to a lesser extent delayed in two plant species, *Trifolium pratense* and *Taraxacum officinale* ssp. *vulgare*, exposed to increasing doses of the herbicide fluroxypyr immediately before opening of the flowers (Boutin et al., 2014). Plants were sprayed with four doses (0, 5, 25 and 100 % of recommended 144 g a.s./ha label rate); time and number of flowers were recorded. The average number of flowers produced by *Trifolium pratense* and *Taraxacum officinale* ssp. *vulgare* was severely impaired at 25 and 100 % label rates and even at 5 % label rate for *T. pratense*.

The time required for plants to flower and recover following herbicidal impacts is likely to play an important role in plant community dynamics. A slight decrease in height or reproduction can put affected, sensitive wild species at a greater disadvantage than more tolerant or insensitive species within the same community. This may undoubtedly be most detrimental for annual species, including rare arable weeds, but may also affect biennials and perennials.

A question arises from the above greenhouse studies: do such effects on flowering and recovery time have an influence on plant populations and communities? An elegant study conducted in the UK started addressing this issue. Gove et al. (2007) conducted an experiment with six woodland plant species. The authors exposed them to the herbicide glyphosate at concentrations ranging between 1 and 25 % of the full application rate, equivalent to spray drift. Plants were six weeks old when they were treated with the herbicide. The study design included a greenhouse and two field components. Plants were grown individually in pots in the greenhouse, and harvested after 10 weeks. The field part included the same species grown in woodlands, but this time harvested after one year. In addition, plants were surveyed in semi-natural habitats adjacent to crop fields. The study found increased mortality, reduced biomass and reduced fecundity in all six species tested, both in greenhouse experiments, where plants were grown separately and exposed when they were six weeks old, and when transplanted into plots in woodland margins. However, plants exposed and treated under field conditions were more affected than plants in the greenhouse. This suggests that short-term greenhouse studies may underestimate toxicity. The field survey revealed that the most sensitive species in the greenhouse was also the least abundant in fields with high agricultural herbicide inputs.

Two additional studies assessed the effects of exposure to the herbicide glyphosate on plant populations and communities (Perry et al., 1996; Strandberg et al., 2012). Both studies were designed to determine the combined effects of glyphosate and fertiliser on experimentally established vegetation mimicking effects on field margins. Perry et al. (1996), who only reported first-year results, found that, although each individual species (three monocots and three dicots) responded differently to the treatments, both fertiliser and glyphosate affected the community significantly. Concurrently, Strandberg et al. (2012) found species-dependent responses to glyphosate, and they showed interactions between nitrogen and glyphosate on species composition, species richness and total biomass. The study by Strandberg et al. (2012) was established in 2001 and included effects following nine years of exposure to 12 treatments (combinations of low dosages ranging from 0 % to 25 % label rate of 1 440 g a.s./ha) of glyphosate and nitrogen on sown vegetation comprising 31 species. Significant effects of glyphosate concentrations equivalent to those measured in spray drift (1–25 % of full application rate) were shown, and the responses of the vegetation were affected by the application of fertiliser. Efficacy studies of herbicides also indicate that the herbicide sensitivity of the different weed species was influenced by N level (Cathcart et al., 2004).

Additional long-term studies were conducted in which herbicides were applied annually for two to four years. Marrs et al. (1991a, b) and Marrs and Frost (1997) showed that effects of herbicides (glyphosate, MCPA and mecoprop) on yield and reproduction were more pronounced in the second year for most species. However, the growth of some species was enhanced by herbicide, which may be due to a shielding effect or a competitive advantage of less sensitive species. It is very likely that species located within agroecosystems will repeatedly be at the receiving end of herbicide droplet drift and that long-term consequences of plant communities is to be expected. Current regulatory guidelines do not take into account effects of recurrent herbicide use.

2.11. Indirect effects on other trophic levels

Plants are the primary producers and the foundation of terrestrial ecosystems. The link between terrestrial primary producers, biodiversity and interaction networks is well recognised (Hooper et al., 2005; Proulx et al., 2010; Pocock et al., 2012). Declines and modifications such as those documented with respect to herbicide use would subsequently have consequences through food web interactions, negatively affecting other trophic levels (e.g. arthropods, birds, mammals) through loss of food and habitat resources (Potts, 1980; Sotherton et al., 1988; Biesmeijer et al., 2006). For instance, seeds are

produced by plants in large numbers. However, most of them do not germinate to become seed-producing plants, but rather are consumed by organisms such as fungi, other microorganisms, invertebrates or vertebrates. One of the major problems caused by the lack of seeds is the loss of a food source.

A diminution of plant flowering and fruiting could have a dramatic effect locally on pollinators and later on fruit- or seed-eating animals. Studies have revealed that, as a result of herbicide use, the decline of cover and diversity of flowering plant species in crop fields and field margins has subsequently reduced resources available for flower-visiting insects and other arthropods (Lagerlöf et al., 1992; Longley and Sotherton, 1997; Holzschuh et al., 2007). Likewise, abundant floral diversity was found to be the prevailing factor related to high Lepidopteran diversity in farmland habitats (Boutin et al., 2011, and references therein).

Indirect effects of herbicides can be very subtle. Caterpillars feeding on *Ranunculus acris* treated with 3 % application rate of a sulfonylurea herbicide (Atlantis® WG) showed statistically significantly lower weights than caterpillars feeding on untreated control plants Schmitz et al.(2013b). Since the herbicide showed no direct toxicity towards the caterpillars, the results indicate a reduced host plant quality of *R. acris* possibly caused by defence components produced in the plants following the herbicide application.

By reducing plant species diversity and abundance in the crop fields and their margins as well as in adjacent habitats, the diversity and abundance of invertebrates is reduced, which in turn will affect other trophic levels. One of the best documented studies of the consequences of alterations of plant species composition and habitat quality was performed in Britain. The grey partridge (*Perdix perdix*) has been surveyed since 1933, and it was found that numbers declined by 80 % between 1952 and the mid-1980s (Sotherton et al., 1988). The plummeting number of grey partridges in agricultural land was attributed to declining chick survival early in the season owing to weed removal by herbicides, causing a shortage of insects at this very crucial period of the year (Potts, 1980). Modern agricultural practices (removal of hedgerows, reduction of margins) with the subsequent alteration of preferred nesting sites was also a contributing factor.

Mammal species often coexist by occupying structurally distinct habitats and utilising different food sources. Potential indirect ecological effect of herbicide use on mammals is better depicted through a case of herbicide use for the eradication of pocket gophers (Tietjen et al., 1967). This species (*Thomomys talpoides*) survives by consuming broadleaved forbs. The herbicide 2,4-D sprayed over pasture land in Colorado caused the vegetation to shift from a mainly broadleaved plant community to a land dominated by grasses. As a consequence of the removal of their preferred food, the animals had to move to an unsprayed area. The subsequent effect on reproduction and survival of the population was not investigated but it was shown in another study that reduced forage abundance had an impact on survival rates of pocket gophers (Hull, 1971). Many examples have been documented which link changes in diversity of plants with changes in small mammal communities in forested areas (e.g. Sullivan et al., 1998). However, studies on the effects of herbicides on mammals in agroecosystems are scarce. Tew et al. (1992) studied wood mice activities in field margins sprayed or unsprayed with herbicides. It was found that wood mice spent significantly more time in unsprayed plots where food availability was high (plants and invertebrates) than in sprayed plots. Furthermore, observations suggested that the time spent in the food-rich patches was largely spent feeding. It seems likely that alterations of plant communities through herbicide use would be detrimental to other small mammals and their predators and, eventually, would have a noticeable effect on the composition of mammals and other vertebrate communities. Indirect effects on higher trophic levels may not only act via the alteration of the food web. In particular, small mammals using crop fields as a feeding habitat may be vulnerable to the alteration of habitat quality owing to a reduction in ground cover which increases their risk of predation (Jahn et al., 2014).

Population-relevant impacts owing to indirect effects of herbicide use may cause a shortage of food availability and suitable habitats. The reduction of weeds and non-weeds in the crop area owing to

pesticide application may lead to a situation where the protection of species at higher trophic levels, such as arthropods, birds, mammals or amphibians, is seriously hampered owing to the fact that the scarcity of resources in the crop area cannot be sufficiently compensated for by non-crop areas.

2.12. Protection goals and agricultural and management practices

The choice of appropriate risk management options clearly depends on the definition of the SPG. Whereas adequate protection of the plant communities in the off-field area can be achieved by exposure-reducing risk mitigation measures, this is mostly not an appropriate option for the management of the risk to non-target species in the in-crop area. This is because reducing in-field exposure (i.e. reducing the application rate) to an acceptable level would challenge the intended plant protection, especially in the case of herbicides. Therefore, risk mitigation measures for in-crop SPGs should aim to compensate for unavoidable effects rather than reduce exposure. Indeed, indirect effects both in field and off field owing to PPP use need to be compensated for by appropriate measures (MAGPie risk management workshop, 2013: mitigating the risks of PPPs in the environment), including describing compensation measures as an option for managing in-field effects of PPP.

Risk mitigation measures implemented at the EU level in the authorisation procedure of PPPs focused on the reduction of exposure of the off-field area. The only measures currently accepted by all EU Member States are non-spray areas at the edge of the field by in-field buffer zones to adjacent off-field areas. The focus of a non-spray area in-field (buffer zone) is primarily on the reduction of drift and run-off entries from treated fields into adjacent off-field areas (see sections 2.3, 2.4.2 and 5).

Many EU Member States also apply drift-reducing application techniques such as low drift nozzles or directed applications in order to reduce the exposure via spray drift and dust drift outside the field of application (see sections 2.3 and 5). It should be noted that drift-reducing techniques do not affect long-distance transport and the subsequent deposition of volatile substances far from the application area. Therefore, other risk management options such as the restriction of the application of highly volatile PPPs to certain environmental conditions or changes in the PPP composition/formulation (e.g. micro-encapsulation) have to be considered on a case-by-case basis.

Additional measures exist in different Member States to mitigate or compensate risk owing to direct and indirect effects of PPPs both in field and off field (for details refer to DEFRA, 2004; Bright et al., 2008; Jahn et al., 2014). For example, some of these management measures are suggested by integrated pest management (IPM) (Prokopy, 2003; Ehler, 2006; Reichenbergeret et al., 2007; van Eerdt, 2014). IPM is mainly composed of exposure mitigation measures for PPPs. Several measures aim at reducing exposure such as using alternative PPP formulations, patch spraying, restriction of application of PPP in ecological hot spots (nesting sites, burrows, see Jahn et al., 2014), and alternative methods of cultivation or use such as low pesticide-input agriculture (e.g. mechanical weed control). Other measures have the primary aim of compensating for in-crop effects on higher trophic levels by providing alternative in-field areas with improved food availability that also serve as alternative habitats (e.g. conservation headlands; creation of areas with sparsely sown cereal crops and restriction of application of PPP; creation of flowering areas or strips; keeping over-wintered stubble with self-greening and as appropriate with maintenance measures, and whole-field set-aside). If designed as a buffer zone between in- and off-field areas, these compensation areas could additionally contribute to the exposure reduction for off-field areas.

There exists a large variety of options for the mitigation or compensation of inevitable effects of PPPs on arable plant species of high conservation value and biodiversity of the agroecosystem in general owing to indirect effects on higher trophic levels. Whereas most options mentioned above can be expected to improve the food provision to higher trophic levels, the appropriate options for the conservation of the arable flora are mostly those where cultivation of the area is still retained. The concrete risk management concept (including the choice of the adequate risk management measures and their combination) needs to be established on a national level, reflecting ecological and agricultural conditions such as the availability of drift-reducing application techniques or national

policies for implementation of management and mitigation options, e.g. obligatory vegetated buffer strips of a certain width.

2.13. Conclusion and highlights

- Important ecosystem services are provided by NTTPs in terms of nutrient cycling and water regulation, and they are supporters of food webs, genetic resources and aesthetic values, as well as biodiversity.
- Specific protection goals and test endpoints should be driven by ecosystem services provided by NTTPs defined for in- and off-field areas.
- Vulnerable species and growth stages have been defined and can include rare arable weeds and sensitive species underrepresented in testing schemes at phenological stages not usually measured in risk assessment.
- Many arable weeds have become rare due to agricultural intensification in several European countries, including Denmark, Germany, Spain, France, the Netherlands, Sweden, the UK and Turkey. This is also true for other wild species, i.e. all non-target plants.
- Herbicide use has been identified as one the main factors for this decline.
- Rare arable weeds are usually annual species that need regular soil disturbance and are preferably found in crop edges of conventional farming as well as in field centre and edges of organic fields.
- Management practices that favour rare arable weeds have been identified, e.g. uncropped tilled field edges with no herbicide spray.
- Effects of herbicides on plant reproduction can be more pronounced than effects on vegetative parts.
- Effects on reproduction can be especially detrimental for annual species, including rare arable weeds.
- Long-term effects and the ability of plants to recover when sprayed at sublethal doses of herbicides can be significant.
- Plant populations and communities can be affected by sublethal doses of herbicides through effects on vegetative biomass, height and seed production.
- Herbicides can affect the species pool of a plant community by reducing reproduction and subsequently the seed bank.
- Soil seed banks are an important component of vegetation dynamics in ecosystems and constitute a reserve of biodiversity, especially for short-lived species.
- Modifications of primary producers and seed reduction can affect other trophic levels in terrestrial ecosystems and result in indirect effects on biodiversity.
- A variety of risk mitigation options for the in- and off-field risk are available.
- The implementation of in-field buffer strips as mitigation measures may have the side effect that farmers will increase the size of their field in order to minimise the total area of field surface to be used for buffer strips. For the sake of conserving overall biodiversity in agricultural landscapes, this should be avoided, particularly if the relative area of off-field habitats is relatively small in the landscape of concern. In such cases, areas for in-field compensation measures that are proportional to field size should be implemented, e.g. wider in-field buffer strips if the size of the fields is larger.

3. Selection of species

3.1. Crop and wild species

Previous research demonstrated that crop species do not consistently differ from wild plant species in herbicide sensitivity (Boutin and Rogers, 2000; McKelvey et al., 2002; Clark et al., 2004; Olszyk et al., 2006). Olszyk et al. (2008) found that wild species usually fell within the range of sensitivities of crop species but were more variable in sensitivity than standard crop species, whereas Boutin and Rogers (2000) noted that risk assessments based on crop data alone were not necessarily protective of sensitive wild species. This is in agreement with Boutin et al. (2004), who indicated that wild species (15 species tested with six herbicides) were often more sensitive than the crop species most commonly used in regulatory testing as per the US Environmental Protection Agency (EPA) online database. Strandberg et al. (2012) re-analysed two datasets, i.e. the 15 non-crop species from Boutin et al. (2004) and crop species from the PHYTOTOX and ECOTOX (http://www.epa.gov/ecotox/) databases. Although they found a consistent trend towards crops being generally less sensitive than non-crop species, it was only significant in the case of glyphosate. In a new experiment, Strandberg et al. (2012) tested 10 crop species and two non-crop species, *Centaurea cyanus* and *Papaver rhoeas*, common to the experiment by Boutin et al. (2004). They first compared the ER_{50} of the two non-crop species common to the two studies. They found that the ER_{50} was similar for the herbicides bromoxynil and glyphosate, while there was a large discrepancy in the ER_{50} for metsulfuron methyl, with plants exhibiting higher sensitivity in the study by Strandberg et al. (2012). For metsulfuron methyl, it was also revealed that crops were less sensitive than non-crop species, whereas there was no difference in the case of bromoxynil and glyphosate. Overall, species sensitivity was dependent on the efficacy spectrum of the herbicide and whether or not the test species was a monocot or a dicot species. Furthermore, results indicated that variation in test conditions may be more important for the previously observed differences in sensitivity of crops and non-target species than whether it is a crop or a non-target-species. Most crops are annuals but no clear tendency was observed in these studies between annual and perennial species (see also section 3.2). However, the datasets were limited.

Most studies presented above included data from several sources with varying experimental methods and pesticide formulations. This observation led the way to more elaborate experiments conducted using crop and wild species tested under the same conditions and same time periods. White and Boutin (2007) investigated the sensitivity of crops and taxonomically or morphologically related wild plants to herbicides. Nine crop and nine non-crop species were paired and tested with five formulated herbicides: atrazine, MCPA, imazethapyr, bentazone and glyphosate. Plants were sprayed at the two- to six-leaf stage and harvested 28 days after spray. Statistical tests conducted for each herbicide comparing the ER_{25} of crop and wild plant species showed that there was no significant difference in herbicide sensitivity between species types for any of the herbicides included in the study. Carpenter and Boutin (2010) used 10 crop species (four monocots from two families; six dicots from five families) and 10 wild species (four monocots from two families; six dicots from six families) following the guidelines established by the United States Environmental Protection Agency (US EPA, 1996, revised in 2012). Wild monocot and dicot plant species were selected for their ease to grow and seed availability but with no other particular relatedness to the selected crops. The study was conducted under greenhouse conditions following current regulatory guidelines whereby plants were sprayed with the formulated herbicide containing glufosinate ammonium at the three to six true leaf stage and harvested 21 days after spray. There was no clear difference found between crops and wild plant species in terms of sensitivity (ER_{50} values) following glufosinate ammonium exposure, although the herbicide caused significant decreases in total aboveground biomass in the majority of species tested when harvested three weeks after spray. Schmitz et al. (2013b) presented two studies where crop and non-crop species were compared with glyphosate and a mixture of sulfonylureas. They found that most non-crop species were more sensitive than crops, but not in all cases.

Results of the above experiments have other implications for regulatory testing. White and Boutin (2007) showed that crops in which a specific herbicide is intended for use (based on the herbicide labels), along with their wild relative, have a higher tolerance to that herbicide than other species. It

seems that current pesticide registration guidelines are too rigid in terms of species selection (US EPA, 2012), requiring maize and soybean, as well as three other monocotyledonous species, to be included in every herbicide risk assessment regardless of known tolerances to certain herbicides. Likewise, the EPPO guidelines (226/1, 2003) are very specific in their plant requirement (see above). Ideally, guidelines would allow the selection of species included to be tailored to a specific herbicide or even region/habitat in which the herbicide may be utilised to better protect species of conservation interest as well as agronomically important species (Pfleeger et al., 2006). The use of native species would also aid pre- and post-registration monitoring to determine the success of restrictions imposed on pesticide labels in protecting non-target habitats (Olszyk et al., 2006).

In the case of the six herbicides tested in two of the experiments described above (White and Boutin, 2007; Carpenter and Boutin, 2010), crops would appear to be suitable surrogates for wild species when plants are tested at the juvenile stage. It was not the case for the experiment in Schmitz et al. (2013b). It has been shown in numerous studies that species sensitivity varies considerably with the herbicide tested and that no one plant species is consistently the most or the least sensitive (Fletcher et al., 1985; Marrs et al., 1989; Pestemer and Zwerger, 1999; Boutin et al., 2004; Clark et al., 2004; Strandberg et al., 2012). However, non-crop species were better suited for testing herbicide effects on reproduction (Carpenter and Boutin, 2010; Carpenter et al., 2013). Therefore, if conservation of wild species is the primary intention, then ecologically relevant test species should be favoured in phytotoxicity testing, alongside the agronomically significant species for non-target crop protection.

3.1.1. Germinability of crop and wild species

Historically, species selected for inclusion in phytotoxicity testing were crop species because they often have large seeds with no particular requirements for germination, are readily available from seed companies and produce consistent and reliable rates of germination. However, non-crop species have been used time and again in experimental studies for different purposes, including phytotoxicity studies (OECD, 2006a, b; US EPA, 2012a), and many species have shown to be easy to manipulate and to yield uniform germination. White et al. (2009) tested the germination rate and requirements of 29 terrestrial and wetland plant species that are regularly found in marginal habitats near cropland. Many were herbaceous perennials and biennial species, contrary to most crops that are annual species. They found that 23 of the tested species reached 70 % germination, while an additional six species reached 50 % germination with minimal stratification and light requirements. The 70 % germination was considered the acceptable level for the seedling emergence test (OECD, 2006a). Most species reached 70 % germination in less than 14 days. Regulatory guidelines (OECD, 2006a,b; US EPA, 2012) include a list of 53 herbaceous non-crop plant species suitable for testing based on a thorough literature review. Of the species tested in White et al. (2009), 12 were part of the proposed list of recommended species in the Organisation for Economic Co-operation and Development (OECD) guidelines (2006), and 10 of them achieved germination and time to germination well within the criteria deemed acceptable for phytotoxicity testing (OECD, 2006a, US EPA, 2012a). Boutin et al. (2010) later tested the germination of different ecotypes (plants originating from different areas of the world) for eight non-crop species. A few ecotypes of the same species differed in their seed size, percentage germination and germination requirements, illustrating the need for further studies to elucidate the cause of the discrepancies. Although numerous phytotoxicity studies have successfully been conducted using non-crop plants (Boutin et al., 2000, 2004; Riemens et al., 2008, 2009; Carpenter and Boutin, 2010; Strandberg et al., 2012; Carpenter et al., 2013), in most cases, germination characteristics were not tested. Nevertheless, they demonstrated that using non-crop species in greenhouse testing is straightforward.

Two other studies have been conducted with the objective of testing germination of non-crop plants. As early as 1993, Cole et al. tested the germinability of 22 non-crop and seven crop species. In two separate experiments, seeds were sown to 2 cm or 1 cm depth. All crop species and only six non-crop species reached the 70 % threshold overall. Most non-crop species need light to germinate and sometimes some stratification. It is believed that the conditions provided to non-crops for germination were inadequate. Similarly, Pallet et al. (2007) tested the emergence six non-crop species

recommended in the OECD (2006a,b) guidelines. Unfortunately, sowing depths ranged from 2 mm to 10 mm; therefore, except for *Ipomoea hederacea* (L.), little germination occurred.

3.2. Annual and perennial species

In natural and semi-natural habitats, perennial species normally dominate, although a number of annual and biennial species may also be found, including some rare arable weeds. Among the studies that tested the sensitivity of NTTPs to herbicides, a number of studies have included annual as well as perennial species. The study by Strandberg et al. (2012) compared the sensitivity of three annual species and three taxonomically closely related perennial species to three herbicides (glyphosate, metsulfuron-methyl and mecoprop-P). They found no systematic differences in sensitivity among annuals and perennials. However, annual/biennial species that need to produce viable seeds and go through the sensitive early growth stages every year or relatively often are in general more vulnerable to herbicides than other species. A study of succession of experimentally established grassland exposed to low dosages of glyphosate (0–360 g a.i./ha) reinforces this as several biennial species which were common had almost disappeared over a five-year period (Strandberg et al., 2012). Several studies conducted in Canada and Denmark have shown that there is no significant difference between the sensitivity of short- and long-lived species in terms of intrinsic sensitivity (Boutin et al., 2004; White et al., 2007; Carpenter and Boutin, 2010; Boutin et al., 2012; Carpenter et al., 2013). However, plants with different lifespans would possibly show differences in their recovery potential following a stress event such as pesticide exposure.

3.3. Crop varieties and wild ecotypes

A potential area of weakness in current pesticide registration guidelines is the inclusion of data for only one crop variety (cultivar) for each of the tested species. Any given crop species has many different cultivars, and it has been known for some time that levels of herbicide tolerance differ among cultivars as this is verified in the context of crop margins of safety when pesticides are registered. This has also been documented in the scientific literature for maize (*Zea mays*) (Keifer 1989; Rowe et al., 1990; Burton et al., 1994), soybean (*Glycine max*) (de Weese et al., 1989; Hulting et al., 2001, Wax et al., 2006), pumpkins (*Cucurbita* spp.) (Harrison and Keinath, 2003), potatoes (*Solanum tuberosum*) (Freisen and Wall, 1984; Arsenault and Ivany, 2001), cotton (*Gossypium hirsutum*) (Abernathy et al., 1979), dry edible beans (*Phaseolus vulgaris*) (Urwin et al., 1996) and cabbage (*Brassica oleracea*) (Hopen et al., 1993) and is likely for many other crop species.

Historically, registrants have been reluctant to use wild species, based on allegations that a large variability may exist among different ecotypes (or populations adapted to particular sets of environmental conditions). Nevertheless, a list of wild species that has been successfully used in toxicity testing is included in the OECD guidelines (2006a,b). Furthermore, many wild plant species have been successfully used in phytotoxicity studies in recent years (Brown and Farmer, 1991; Cole et al., 1993; Kjær, 1994; Breeze et al., 1999; Boutin et al., 2000, 2004, 2010; Blackburn and Boutin, 2003; Olszyk et al., 2006; White and Boutin, 2007; Riemens et al., 2008, 2009; White et al., 2009; Carpenter and Boutin, 2010; Strandberg et al., 2012; Carpenter et al., 2013). Many good candidate wild plant species that can be used in phytotoxicity testing are circumpolar in their distribution or have spread to different regions through human introductions; thus, several ecotypes exist among wild populations (Aude et al., 2003; Bleeker et al., 2007; Guo et al., 2009; Boutin et al., 2010).

Two experiments were conducted to measure the variability in herbicide sensitivity among cultivars of the same crop species, and to assess the variability in phytotoxicity response among ecotypes (plant populations) of wild species originating from different parts of the world. In White and Boutin (2007), crop (lettuce, *Lactuca sativa*; radish, *Raphanus sativus*; tomato, *Solanun lycopersicon*; maize, *Zea mays*; and onion, *Allium cepa*) cultivars ($n = 5$ cultivars for all crops, except $n = 3$ for onion) were exposed to the formulated herbicides containing atrazine, imazethapyr and MCPA in separate experiments under greenhouse conditions. Plants were sprayed at the two- to six-leaf stage and harvest of the aboveground biomass was carried out 28 days after spray. Results showed a significant difference between cultivars for all species. The study also revealed that cultivar sensitivity was not

only variable but also herbicide dependent. It was concluded that the phytotoxicity range of any given herbicide may differ depending on whether the cultivar chosen for inclusion in the toxicity test is tolerant or sensitive to a specific herbicide. These results revealed that the selection of the cultivar could alter the outcome of the risk assessment conducted during the regulatory process and introduce uncertainty in the assessments.

The variability among ecotypes of wild species was tested under greenhouse conditions with eight species using two to four ecotypes per species from different areas of North America and Europe (Boutin et al., 2010). Plants were tested with a formulation of atrazine and glyphosate at the three- to five-leaf stage, and the aboveground biomass was harvested 28 days after spray. In spite of the fact that the dose–response curves of ecotypes for a given species were remarkably similar, with the exception of *Prunella vulgaris* tested with atrazine, results revealed a significant difference for six out of eight species and five out of eight species for atrazine and glyphosate, respectively. Herbicide sensitivity (ER_{25}) among ecotypes of the same species varied from a factor of less than two in eight cases (out of 16) to more than one order of magnitude in one case with atrazine. As for crop cultivar (White and Boutin, 2007), the present study demonstrated that conclusions regarding the phytotoxicity of any given herbicide may differ depending on the ecotypes chosen for inclusion in risk assessment. Several characteristics of the studied ecotypes were found to differ including seed size, percentage germination, germination requirements and growth patterns; however, these differences were not found to be related to any pattern of sensitivity observed among the species tested. These findings agree with previous research, in which disparity in herbicide susceptibility among ecotypes has been confirmed for several weed species (DeGennaro and Weller, 1984; Klingaman and Oliver, 1996; Noldin et al., 1999).

3.4. Testing woody species

In regulatory testing, species selection is usually limited to a narrow taxonomic range of usually only short-lived crop species. This raises the question of whether or not other types of plants, such as woody species, will be protected. A study was undertaken to compare the sensitivity of woody and crop species with the herbicide PAR III® (Boutin et al., 2012). PAR III® is a formulated herbicide comprising mecoprop (61.6 %), 2,4-D (32.5 %) and dicamba (5.8 %). Plants were grown under greenhouse conditions and sprayed at the four- to six-leaf stage and the aboveground biomass material was harvested 28 days after spray. In all cases, there was one plant per pot. The three grass crop species (*Lolium perenne*, *Zea mays* and *Avena sativa*) were very resistant to the herbicide with little effect at 100 % of the recommended label rate (1 848 g a.s./ha) (Figure 1). Results also revealed that this herbicide mixture was not very toxic to five of the seven woody species when sprayed at the young vegetative stage. For example, *Rhamnus cathartica* and *R. frangula* were not sensitive to the herbicide PAR III®. They are both introduced species in eastern Canada that are considered invasive in forested areas, inhibiting the establishment and growth of native understory herbaceous and other woody native plants. These species would be at an advantage if spray drift reached a forested community adjacent to a spray area (e.g. woody hedgerow) or if overspray occurred. Conversely, elm (*Ulmus americana*) and poplar (*Populus grandidentata*) trees were very sensitive and would be affected at less than 10 % of the recommended label rate. However, overall, the species sensitivity distribution using the 13 crop and 7 woody species of this experiment revealed that current tests with crop species may be sufficient for the assessment of woody species with this herbicide when plants are tested at the seedling stage (see Fig. 1 of Boutin et al., 2012). Further studies are needed with additional herbicides tested at various phenological stages to confirm these results.

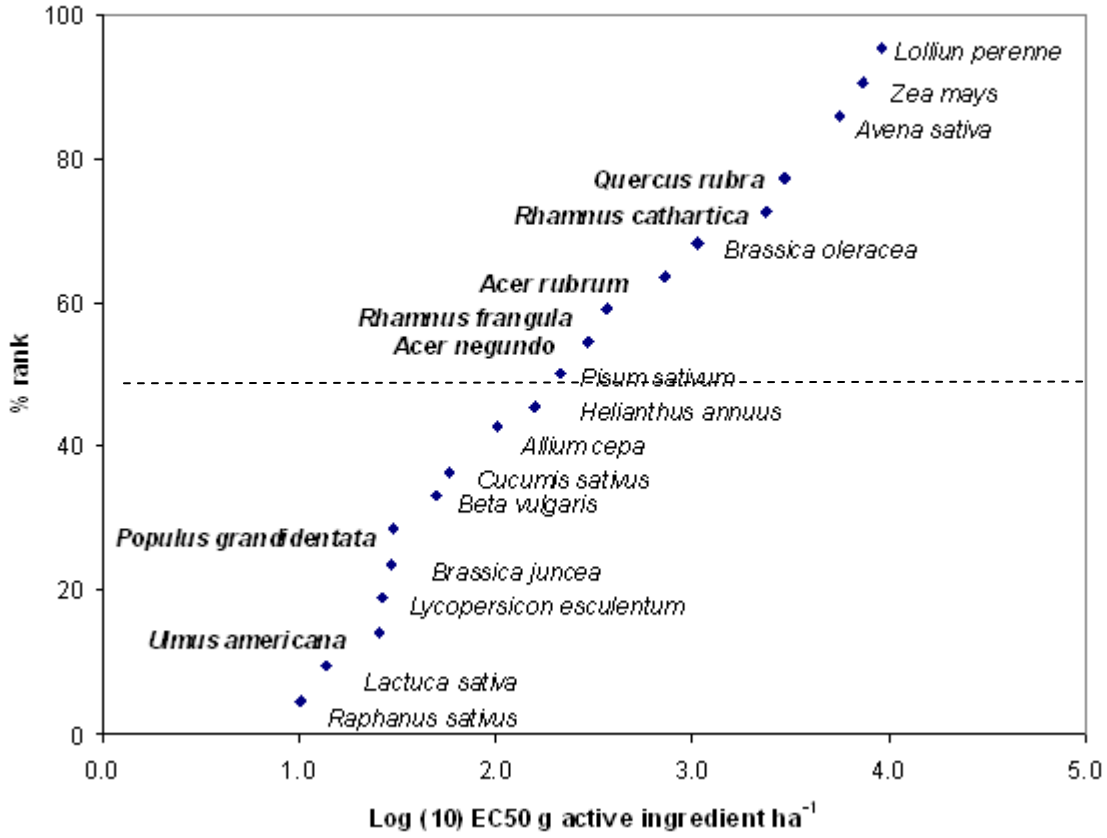


Figure 1: Ranks of species sensitivities plotted for 20 species tested with the herbicide PAR III®. The seven woody species, in bold, were found in a protected area of Ottawa, Ontario, Canada, and were compared with crop species normally used for risk assessment. Species below the dashed line were affected at doses 10 % below the recommended label rate of 1 848 g active substance/ha (from Boutin et al., 2012).

Marshall (1989a) tested four shrub species with 15 herbicides and three growth regulators. Plants from a commercial nursery were two years old with heights ranging from 30 to 40 cm at the start of the experiment. They were maintained in a glasshouse subjected to ambient conditions until spray time in the spring. Two doses were used: half of and the full recommended label rate for soybean. Plants were assessed visually 18 weeks later and the aboveground material was harvested one year after spray. All herbicides caused some effects on one or all shrub species tested. Five herbicides (mecoprop, fluroxypyr, chlorsulfuron, metsulfuron-methyl and glyphosate) caused significant damage to most species. In a field experiment conducted in the USA, Fletcher et al. (1993) found that cherry trees sprayed at doses as low as 0.2 % of the field application rate of the sulfonylurea chlorsulfuron showed a significant reduction in the production of fruits, with almost no observable damage to vegetative parts. Bhatti et al. (1995) also noted the high sensitivity of cherry trees but with variable effects on the reproduction of various cultivars. In these studies, effects owing to a one-time application of herbicides were evaluated.

The long-term impact of recurrent herbicide applications on trees and shrubs growing adjacent to crop fields has never been assessed, that is small amounts of herbicides reaching woody plants causing

sublethal effects several times a year and repeatedly over several years remain unknown. However, an experiment conducted under normal field conditions in Denmark aimed to evaluate effects after one year following a one-time application of the sulfonylurea metsulfuron methyl. It was found that the reproductive endpoints (e.g. green and mature hawthorn, *Crataegus monogyna* Jacq., berries) were severely affected by average spray drift concentrations at higher than 2.5 % of the label rate of metsulfuron methyl, and that the effect was still observed one year after the spray event (Kjær et al., 2006a, b). Reducing flowering and fruiting of not only cultivated woody plants but also wild woody species could have dramatic long-term effects on other trophic levels. As with herbaceous species (Carpenter and Boutin, 2010; Strandberg et al., 2012; Carpenter et al., 2013), woody species may be especially sensitive at the blooming stage, and this should be investigated further.

3.5. Testing with cryptogams

Fern species and other cryptogams are never used in phytotoxicity testing for regulatory purposes, despite the fact that many species constitute important components of some of the moist and shaded habitats adjacent to crop fields in North America and Europe. Boutin et al. (2012) presented an experiment conducted with two fern species, *Onoclea sensibilis* and *Dennstaedtia punctilobula*. *O. sensibilis* is a common fern species found in a wide range of habitats, including ditches and hedgerows adjacent to cropped field margins, whereas *D. punctilobula* grows mostly in forested areas of eastern Canada. The objective was to assess the sensitivity of the two fern species to two formulated herbicides containing glyphosate (12 % and 28 % label rate or 213.6 g and 498.4 g a.i./ha) or metsulfuron methyl (1 % label rate or 0.045 g a.i./ha). Plants were grown from spores and were tested at the early sporophyte stage. A non-ionic surfactant Agral 90 containing nonylphenoxy polyethyoxyethanol was added to metsulfuron methyl concentration as recommended on the label. Ferns were grown under greenhouse conditions and harvested four weeks after spray and aboveground dry biomass was obtained. Results demonstrated a high sensitivity to metsulfuron methyl and to a lesser extent to glyphosate of both species when tested at the early sporophyte stage. Glyphosate caused a significant reduction in biomass of the two fern species at 12 % and 28 % of the label rate, whereas metsulfuron methyl significantly decreased biomass at 1 % label rate.

In another study carried out in North America, but this time under field conditions, the sensitivity of two herbicides on cryptogam species was clearly demonstrated (Newmaster and Bell, 2002). The experimental work was conducted in Northern Ontario, Canada, with two formulated herbicides, triclopyr and glyphosate, which are commonly used in forestry for conifer release. Effects of the two herbicides were assessed in separate experiments on pteridophytes, bryophytes and lichens. The dose usually applied for each herbicide was sprayed via helicopter; however, the dose actually reaching the cryptogam (understory) plants is not known. Botanical surveys were conducted before and every year after treatments for five years. The two herbicide treatments had a significant initial effect on species richness, species abundance and diversity of all three groups. Although signs of recovery occurred, it was still not totally achieved five years after treatment, partly owing to the indirect effect mediated through reduced vegetation cover and thus an increased exposure to the sun and wind which produced a dry microclimate not hospitable for most cryptograms. Mosses seemed particularly sensitive to herbicide treatments and changes of microclimatic conditions. It would be relevant to confirm these results with phytotoxicity tests conducted in greenhouses.

In Europe, a few studies have demonstrated the effect of the herbicide asulam to the fern *Pteridium aquilinum*, which is considered a problematic species in upland areas (Marrs et al., 1992; Le Duc et al., 2000). A few other studies have examined the effects of asulam on other desirable fern species. Rowntree and Sheffield (2005) investigated effects on eight fern species tested at the mature sporophyte stage with three doses including full application rate (4.4 kg a.s./ha) and two doses corresponding to 10 m (0.44 kg a.s./ha) and 50 m (0.02 kg a.s./ha) downwind drift from aerial spray. Damage was assessed over two seasons. Maximum damage occurred one year after treatment with limited signs of recovery only seen by the second season. Four of the eight species were affected by the high and medium doses, suggesting that a 50-m buffer zone would be sufficient to protect sensitive ferns. Three of the eight species were more sensitive than the flowering species *Rumex acetosa* tested

at the same time. The fern species tested were exposed when fully mature and thus when they were likely to not be at their most sensitive growth stage.

Rowntree et al. (2003) tested the sensitivity of the mature gametophyte of 18 moss species to asulam. They showed that concentrations as low as 0.05 g a.s./L inhibited the growth of the most sensitive species tested, and the growth of 14 of the species was significantly inhibited at concentrations of 1 g a.s./L. Growth and development of the early life stage (protonema) of three moss species was significantly affected by continuous exposure to asulam at concentrations of 0.01 g a.s./L and above (Rowntree et al., 2005).

There is a shortage of toxicity studies conducted with lichens. Effects of the herbicide trichloroacetic acid (TCA) on lichens was assessed in Finland (Juuti et al., 1996). TCA accumulated in two lichen species and a negative correlation between levels of TCA and biomass was found for one species. Further tests are required to establish this causal relationship.

3.6. Plant trait-based approach and selection of test species

Plant communities are composed of numerous and varied species. Habitats within agroecosystems have been documented to contain hundreds of species in field margins, hedgerows, ditches and other habitats adjacent to crop fields. It has been advocated that using the trait-based approach is more appropriate for linking plant diversity to ecosystem functioning (or ecosystem processes such as primary production, trophic transfer, nutrient cycling, water dynamics) (Diaz and Cabido, 2001) or environmental stressors (disturbance, presence of contaminants) (De Lange et al., 2009).

Traits are the physiological, morphological and ecological attributes of a species that define their role in ecosystems (Baird et al., 2008). For plants, relevant traits can be leaf characteristics (leaf area, leaf mass per area, hairiness, etc.), seed production and morphology, height, root/shoot ratio or root morphology and mycorrhizal association, life history attributes, among others (Westoby and Wright, 2006; Dorrough and Scroggie, 2008; Comas and Eissenstat, 2009; Bernhardt-Römermann et al., 2011).

Advantages of using the trait-based approach:

- can be more sensitive to stressors such as disturbance, grazing, fire, etc.;
- simplifies description of communities;
- avoids redundancy and takes into account species convergence;
- can use simple and easily accessible attributes;
- better links plant diversity to processes and functions of targeted ecosystems;
- reduces the constraints related to scale and geographic differences;
- facilitates comparisons with different species pools.

Disadvantages of using the trait-based approach:

- traits are often unknown;
- intercorrelation among traits;
- phenotypic plasticity for some plant traits.

Dorrough and Scroggie (2008) have demonstrated that simple information about plant species (lifespan, growth form and origin) can be used to predict effects of stressors in pasture land. They found that native perennial species, ferns and shrubs were more impacted by grazing and the addition of phosphorus than exotic annual grasses and forbs. Other studies had addressed functional responses

to management (see Kahmen and Poschlod, 2008, and references therein). Likewise, Boutin and Jobin (1998) revealed that more short-lived grassy-type plants that were originally introduced and are of weedy propensity were more commonly identified in woody hedgerows and woodlots adjacent to intensively managed crop fields (herbicides and fertilisers) than in the same habitat types abutted to less intensively managed fields. In contrast to herbicide drift, which has been well documented (Holterman et al., 1997; de Snoo and de Wit, 1998; Weisser et al., 2002), there is a dearth of published studies on the level of fertiliser misplacement on habitats bordering crop fields (Rew et al., 1992).

Limited experiments have been undertaken to establish the link between herbicide effect and plant traits. A brief review of the scientific literature showed that the relationship between plant traits (physiological, morphological and ecological) and herbicide efficacy is a complex one. Although multiple traits have been suggested as key in explaining species sensitivity, in most cases, single plant traits were measured. These traits include growth habits, leaf morphology, life span, cuticular wax composition and leaf structure characteristics (for example, the presence of trichomes, and stomata size and location) (Benzing and Burt, 1970; Wyrill and Burnside, 1976; Chachalis et al., 2001a; Huangfu et al., 2009). Herbicide absorption generally is aided by cuticular and stomatal infiltration, while cuticular wax is often seen as an effective barrier to absorption (Chachalis et al., 2001b). Species with high stomata density on their leaves may experience relatively higher levels of herbicide uptake, although the stomata are considered as a minimal route for foliar uptake (Huangfu et al., 2009). In addition, leaf surface structures such as trichomes may affect the ability of the herbicides to adhere to the surface of the leaf, thus further reducing the absorption and efficiency of foliar-applied herbicides (Hess et al., 1974). Multiple studies have indicated that the wax composition may be a significant indication of herbicidal efficacy (Mayeux and Jordan, 1980; Wilkinson, 1980; Wilkinson and Mayeux, 1987; Chachalis et al., 2001a, b).

Characteristics of 33 crop and wild plant species were assessed in greenhouse experiments (Boutin et al., 2012). The objective of this study was to assess the influence of selected plant traits on the efficacy of a foliar-applied herbicide, using Roundup Original® containing glyphosate. Traits under consideration in the experimental study included leaf characteristics and plant growth parameters. Although only marginally significant, there was a trend suggesting that leaf area ratio (LAR) (average leaf area per plant (cm^2)/(average weight per plant (g)) and trichome coverage may play a role in determining herbicide sensitivity. LAR is a measure of the net assimilation rates by plants and of the photosynthetic capacity. The results showed a negative trend between LAR and sensitivity to the herbicide glyphosate. Since glyphosate is a contact herbicide, greater leaf surface area means more points of entry for the herbicide, thus increasing efficiency.

The relationship between herbicidal efficacy and trichome density is complex due to the variety in leaf epidermal trichome forms. They are an especially interesting trait, since they can either decrease or increase herbicidal efficacy. Trichomes increase herbicide tolerance when they hinder the wetting and spreading of herbicide droplets (Hull et al., 1982), create air pockets that inhibit contact between the chemical and the leaf surface and/or cause droplets to shatter or bounce away from the leaf epidermis (Hess et al., 1974). Conversely, they may also decrease plant resistance to herbicides by providing an entry site for foliar-applied herbicides (Benzing and Burt, 1970). The positive trend between trichome coverage and the EC_{25}s in this study (Boutin et al., 2012) suggests that trichomes may prevent herbicides from reaching and penetrating the cuticle, thus minimising potential uptake by the plant.

Arabidopsis thaliana was selected for another series of experiments in order to assess which traits of *A. thaliana* were most important in determining glyphosate efficacy (Cognard, 2013). The plant's small size and rapid life cycle are advantageous for research (about six weeks from germination to mature seed). Furthermore, several mutants of *A. thaliana*, which differ only by one specific trait, are available, making it an ideal species for a trait-based approach study. Several mutants and two wild types were tested with glyphosate. Results showed that total leaf area and trichome density were important variables to consider for herbicide effect on *A. thaliana*.

The above studies were only conducted with the contact herbicide glyphosate. Many herbicides with different modes of entry in plants and with different modes of action are used in agriculture. Further studies are needed to determine how and what traits trigger plant sensitivity to different herbicides. Nevertheless, these studies showed that the trait-based approach is a promising avenue to consider for plant species selection in ecological risk assessment.

3.7. Conclusion and highlights

- Commission Regulations (EU) No 283/2013 and No 284/2013 lay down the data requirements that need to be provided as a basic dataset for the authorisation of active substances and PPPs, respectively. Tests are mostly conducted with crop species and requirements are very rigid.
- Experiments have demonstrated that crops may be suitable surrogates for wild species (herbaceous and woody) when tested at juvenile stage under similar conditions, but not always. Moreover, some woody and herbaceous species are very sensitive when sprayed at the reproductive stage (e.g. with sulfonylurea herbicides).
- It would be relevant to use non-crop species for testing, especially since it has not been clearly demonstrated that crop species exhibit the same sensitivity as non-crop species. Until this is properly investigated, risk assessment will remain hard to defend, since tests are performed on crop species.
- Many non-crop species can germinate readily and uniformly under greenhouse conditions with minimum requirements, and are deemed suitable for phytotoxicity testing.
- Annual and perennial species do not consistently differ in their toxicological sensitivity to herbicide, although it remains untested whether they differ in their recovery potential.
- Disparity in herbicide susceptibility among crop cultivars and wild species ecotypes has been confirmed in a number of studies.
- There is a paucity of data on herbicide effects on ferns, mosses, liverworts, hornworts, horsetails, and lichens or woody species. Limited studies showed that they are quite sensitive and may not be sufficiently protected by current risk assessment.
- Using the plant trait-based approach is a promising avenue for plant species selection in phytotoxicity testing and ensuing ecological risk assessment.

4. Effect assessment

4.1. Existing guidelines and schemes

Two main organisations, i.e. OECD and US EPA, have developed guidelines for testing herbicide phytotoxicity to crops and non-target plants. These guidelines encompass exposure before emergence or at early growth stages and effect assessment over a two- to four-week period.

4.1.1. Organisation for Economic Co-operation and Development guidelines

The OECD guidelines, OECD 208 and OECD 227, were developed and published in 2006 (OECD, 2006a, b). The two tests are conducted under greenhouse conditions with plants grown individually in pots or in monoculture. In all cases, a replicate is defined as a pot; therefore, all plants within the same pot are considered as one replicate. The number of species to be tested is not specified and is left to various jurisdictions to be decided. A list of 32 crop and 52 non-crop species historically used in phytotoxicity testing is provided.

The Seedling Emergence and Seedling Growth Test (OECD guideline 208; OECD, 2006a) assesses the effects on vascular plants. The test substance is either incorporated into the soil or applied to the soil surface. Seeds are sowed at the soil surface (usually non-crop species) or in the soil, and herbicide effects are measured 14–21 days after 50 % emergence of the control group has occurred. The endpoints measured are the number and per cent emergence as well as aboveground biomass and

visual injury compared with controls. The Vegetative Vigour Test (OECD guideline 227; OECD, 2006b) assesses the effects on plants sprayed at the two- to four-leaf stage. Plants are evaluated 21–28 days after treatment. Biomass measurements and visual assessments are taken at the end of the test. In both cases, the per cent inhibition is calculated using the EC_x or the concentration that results in a decrease of the test endpoint relative to the control plants. Usually EC_{25} or EC_{50} are calculated corresponding to a 25 % or 50 % reduction, respectively.

4.1.2. United States Environmental Protection Agency guidelines

Three guidelines were made available to registrants by the US EPA in 2012 (https://www.federalregister.gov/articles/2012/06/27/2012-15540/final-test-guidelines-ocspp-850-series-notice-of-availability). The test guidelines further described below are applicable for the evaluation of the hazards and risks of pesticides and industrial chemicals to terrestrial plants resulting from direct or indirect exposure. In addition to these guidelines (US EPA, 2012a, b, c), a background document (US EPA, 2012d) provides general information and overall guidance on test procedures, equipment, statistical analyses and reporting.

The Ecological Effects Test Guidelines: Seedling Emergence and Seedling Growth (US EPA, 2012a) assesses effects on plants exposed at the seed stage through germination, emergence and early seedling growth. The test substance is applied with a sprayer to the soil surface immediately after the seeds have been planted. The dose–response test makes it possible to calculate the EC_{25} using per cent emergence and growth effects. The end-use PPP is used on a minimum of 10 plant species. The species include six dicots from at least four families, one of which is soybean (*Glycine max*); four monocots from at least two families, one of which is maize (*Zea mays*); and at least one is a root crop. Apart from these three recommended crop species, the use of non-crop species is encouraged for the remaining seven species. Table 2 of the guideline is a list of 52 recommended non-crop species identical to the species list in the OECD guidelines (OECD, 2006a, b). The test should be conducted in individual pots under controlled conditions in growth chambers, greenhouses or in small field plots and run for 14–21 days after 50 % of control plants have emerged. One to six seeds are allowed per pot depending on the species, and each pot constitutes a replicate.

The Ecological Effects Test Guidelines: Early Seedling Growth Toxicity Test (US EPA, 2012b) evaluates the effects of a chemical substance applied to newly germinated terrestrial species under indoor controlled conditions. Seeds are germinated in pots or in plugs (for hydroponic test) and germinating seedlings are exposed via either root or foliar routes. The EC_{10} and EC_{50} are determined using seedling survival and shoot, root or total biomass measured after 14 days. The technical grade of the active substance should be used. As in the previous guideline (US EPA, 2012a), six dicots and four monocots should be tested, including three specified crop species (see above) and seven crop or non-crop species.

The Ecological Effects Test Guidelines: Vegetative Vigor (US EPA, 2012c) assesses effects, survival and growth on plants sprayed with a chemical substance (pesticide end-use product) at the two to four true leaf stage. Plants (one to four per pot depending on the species) are harvested 21–28 days after spray. Survival is measured and the EC_{25} and no observed effect concentration (NOEC) (or EC_{05}) are calculated based on plant height and biomass. As in other guidelines (US EPA, 2012a, b), the test is conducted on 10 crop and/or non-crop terrestrial plant species. The species should include six dicots from at least four families, one of which is soybean (*Glycine max*); four monocots from at least two families, one of which is maize (*Zea mays*); and at least one root crop.

These US EPA guidelines above should be used in conjunction with a background document which provides general information and overall guidance on test procedures, equipment, statistical analyses and reporting (US EPA, 2012d).

Testing germinability of plants has long been established in toxicity testing. Seeds are exposed to the test substance on filter paper or in soil. The seed germinability test with direct exposure of the seeds to

compounds added to filter papers is often thought to be sensitive, but unrealistic high exposure should be kept in mind. For seedling emergence tests used in PPPs, regulation seeds are sown in soil, whereas the test substance is incorporated or applied after sowing. This applies for OECD guideline 208 which records effects on seedling emergence as well as on phytotoxicity and growth parameters 21 days after the emergence of 50 % of the seeds. Such a bioassay examines whether the active substance affects germination in or growth through soil in which it is present. It is used both for ecological risk assessment and to evaluate potential negative effects on succeeding crops grown on the treated field owing to residues persisting in the soil (EPPO Guideline PP 1/207, "Effects on succeeding crops"). The second standard test records effects of juvenile plants exposed in the two- to six-leaf stage 21–28 days after application (vegetative vigour, OECD 227; OECD, 2006b; US EPA, 2012c). For many of the tested products, the seedling emergence test (OECD 208) is less sensitive than the vegetative vigour testing (OECD 227). This might reflect reduced exposure via soil compared with via direct spray deposition. Additionally, depending on PPP properties in combination with plant anatomy, uptake of the test substance by green parts may be favoured, whereas uptake into the seed is hampered. However, there are also cases where the seedling emergence test is more sensitive than vegetative vigour testing. This is the case not only for test substances designed to act as germination inhibitors, but also for some products intended to be used on post-emergence stages.

4.2. Using pre-screening, open literature and efficacy data

Plant pre-screening data pertain to the toxicity of pesticide products to terrestrial vascular plants, routinely generated during the early product development process by registrants or manufacturers to measure efficacy (generated for all pesticides). Further data in this document, referred to as "efficacy dossier data", are generated for herbicides to demonstrate action on weeds as well as effects on succeeding and neighbouring crops. Additional efficacy data may be generated for herbicides at a later stage to verify whether or not a compound effectively acts as intended. Data produced to assess crop margin of safety examine whether a compound affects other crops, and they may be available for all pesticides, since registrants need to check the potential impact of all pesticides on crops. In agreement with data requirements (Commission Regulations (EU) No 283/2013 and No 284/2013 of 1 March 2013) screening data for herbicidal activity are often submitted for pesticides where a priori no herbicidal action is assumed. Data investigating the effects on terrestrial plants are required for all plant protection compounds, since pesticides other than herbicides have been found to be phytotoxic (Thomson, 1985). All these data can be used in pesticide risk assessment.

Valuable information and easily accessible data on a wide range of species on effects of herbicides (and other pesticides) on non-target plants that are constituents of wildlife habitats can be provided using all the open literature, plant pre-screening and efficacy dossier data. The ecological relevance of plants routinely tested during the plant screening for PPPs development (pre-screening data) as well as for PPP registration (efficacy dossier data) demonstrated that many of the species were important to wildlife (Boutin et al., 1995). Plant pre-screening and efficacy data are very valuable, even if tests are conducted only for a limited number of doses and no ER_{50} can be calculated, as they include several families and species and, hence, the general spectrum of activity can be determined for each chemical. Effects at maximum label rate are also important information. In addition, data generated for crop margin of safety can provide important information on the range of sensitivity among crop varieties, which may be important for risk assessment to non-target plants, since crop data are usually submitted (Boutin et al., 2010).

During the development process of each PPP, the effects of a chemical on plants are typically assessed by companies. A primary pre-screening for any herbicidal activity is first performed at one high rate on a newly discovered chemical. Once herbicidal activity has been demonstrated, several rates are used in secondary screening to determine weed control efficacy and crop tolerance. Tertiary screening is used to define more precisely the rates of activity. Small-plot field efficacy trials are performed at the fourth level to determine the exact rates of application, the most effective formulations and the effect of adjuvants. Thus, plant pre-screening and efficacy data comprise tests with terrestrial and/or aquatic vascular plants, including tests performed either in the greenhouse or in the field (i.e. pre-plant

incorporated and pre- and post-emergence trials, from primary screening to field trials). They consist of tests performed with one high concentration as well as tests performed with a range of concentrations. Consequently, the data are useful for establishing the range of plant species sensitivity and possibly for determining dose–response curves, EC_{25}s, EC_{50}s, and low- or no-effect levels.

Screening and efficacy data are not performed under good laboratory practices (GLP). However, it is in the interest of the company to perform the tests well to ensure efficacy in controlling weeds and appropriate crop margin of safety.

Screening data on all tested plant species should be provided to give a broad view on the available information on effects of the plant protection compound on terrestrial plants. It might provide information on the range of sensitivity of plant species. Additionally, it should be used to justify the choice of plant species used for the OECD tests where a minimum of six species is used and for the risk assessment.

Visual assessment rating is used during plant pre-screening and efficacy data collection. Below is an example of the rating used which takes into account aboveground biomass qualitatively and visual effects of plants treated with herbicides compared with controls plants.

Common descriptors are chlorosis (i.e. yellowing of leaves), necrosis (i.e. dead parts), malformation, discolouring and formation of mottled patches.

Table 3: Example of the rating on the basis of biomass and visual assessment

Rank	Biomass compared with controls	General characteristics
0	100 %	Healthy
1	> 100 %	Plants significantly larger than controls (hormesis)
2	Slightly < controls	Minor effects or plants slightly smaller than controls
3	~ 75 %	Mild herbicidal damage but not overly affected
4	> 50 %	Obvious herbicidal damage
5	~ 50 %	Plants stunted, often with significant damage
6	< 50 %	Plants stunted, severe damage
7	~ or < 25 %	Severe damage, unlikely to recover
8	< 10 %, or nearly dead	Severe damage, likely to die
9	Dead	Dead

Table 3 shows the relationship between visual assessment rating and aboveground dry biomass, used in existing guidelines (OECD, 2006a, b; US EPA, 2012a,c,). The correlation between the two endpoints is good ($r^2 = 0.74$). However, large differences were observed between compounds and studies (r^2 for studies separately were 0.02 ($n = 10$), 0.04 ($n = 10$), 0.43 ($n = 16$), 0.72 ($n = 19$), 0.72 ($n = 7$), 0.97 ($n = 5$), 0.97 ($n = 6$)). Three studies included the same herbicide with varying r^2 values of 0.04, 0.43 and 0.97. Biomass was a more sensitive endpoint in 46 cases (63 %) and visual assessment was a more sensitive endpoint in 27 cases (37 %). The higher sensitivity of the aboveground biomass endpoint may be due to the integrative nature of this endpoint, although, in some cases, visual effects such as chlorosis, necrosis and discoloration were noted even with no reduction of biomass. Other times, plants may be stunted and malformed without any biomass decline at the time of harvest.

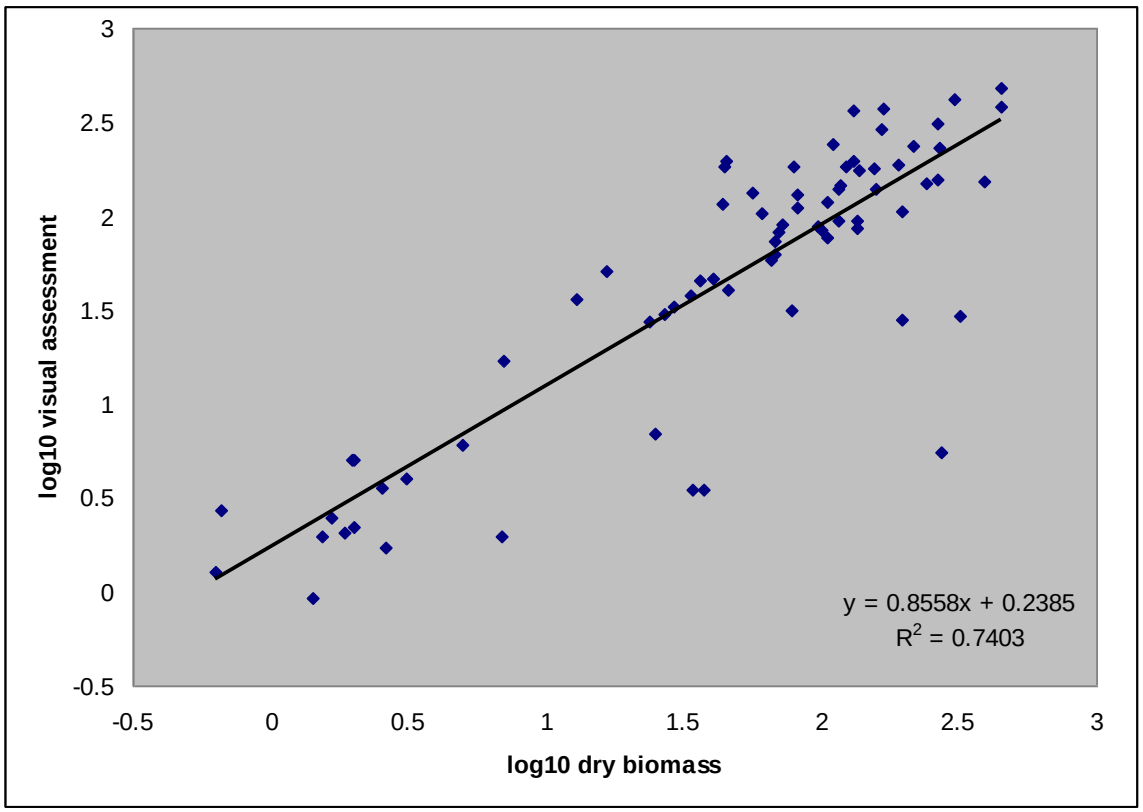


Figure 2: Comparison of effect rates (ERs) calculated using aboveground dry biomass or visual assessment (n = 73). Data were available for seven studies from four herbicides: ER_{20} (two studies), ER_{25} (three studies) and ER_{50} (two studies). (From Boutin, Carpenter, Allison, Parsons, Ellis and Casey, unpublished.)

4.3. Selection of endpoints relative to phenological stage at time of exposure

Four situations can arise when herbicides are sprayed in crop fields and sublethal doses reach non-crop plants through drift into adjacent habitats: (case 1) plants are at the seedling or juvenile stages during spray and the vegetative parts are soon affected; (case 2) plants are at the seedling or juvenile stages during spray and effects on seed production are subsequently observed; (case 3) plants are at the reproductive stage during spray, and flowering or seed production is soon after affected; and (case 4) plants are at the reproductive stage during spray and effects are later observed on the next generation (e.g. seed germination and seedling growth). Thus, species that are at the young seedling stage or species at the reproductive stage may be affected differently and this is likely to be both species and herbicide dependent.

In routine regulatory testing, plants are sprayed during the juvenile period, typically at the two- to six-leaf stage, and effects are recorded 14–28 days after spray, usually when they still are at the vegetative stage (OECD, 2006a, b; US EPA, 2012a, b, c). Indeed, a few greenhouse studies have shown that plants sprayed at the juvenile stage showed more sensitivity than plants sprayed when they were older (eventually at the reproductive stage) when biomass is used as an endpoint, at least in the case of some species (Boutin et al., 2000; Zwerger and Pestemer, 2000; Strandberg et al., 2012). However, when reproductive variables such as seed production are used as endpoints, some studies have shown larger reductions in seed production for plants exposed at the seedling stage than for plants exposed at later stages (e.g. Rinella et al., 2010a,b; Strandberg et al., 2012). This demonstrates that toxicity tests (seedling emergence and vegetative vigour) are not optimised in terms of endpoints (seed production and germinability are not assessed) and species tested (currently focused on crop species).

Appendix A presents data from studies highlighting cases 1 and 2 whereby plants were sprayed at the juvenile stage and effects were recorded at the vegetative (case 1), and subsequently at the reproductive (case 2) stages (Riemens et al., 2008, 2009; Carpenter and Boutin, 2010; Rotchés-Ribalta et al., 2012; Strandberg et al., 2012; Carpenter et al., 2013). ER_{50} and ER_{10} were calculated for the juvenile vegetative endpoint as well as for longer term values on reproduction parameter (seed production). Overall, 41 % of the combinations show a lower vegetative endpoint than for reproduction comparing ER_{50} values. Moreover, 40 % provide a lower vegetative endpoint than for reproduction comparing ER_{10} values. The ER_{10} reproductive endpoint is lower than the ER_{50} vegetative endpoint in 98 % of the cases observed. All outcomes are presented in Appendix A, Table A1. The ER_{50} and ER_{10} were calculated for the vegetative and reproductive endpoints allowing the computation of an extrapolation factor. Although the extrapolation factor is built on a restricted number of studies, these studies are of a high quality. Although effects on seed production have received most of the attention until now, other endpoints are relevant too; for example the effects on plant flowering and the production of nectar and pollen is highly relevant for species that feed on flowers such as pollinators. Many more studies are needed to fully investigate the effects on reproductive endpoints, effects on the next generation and long-term effects of plants living in native plant communities. The extrapolation factor may need to be revised in the light of new information available in the future.

Appendix B summarises studies featuring case 3 demonstrating the effects on reproduction when plants were sprayed during the reproductive stage. Data were extracted from various scientific articles available primarily for sulfonylurea herbicides, which are known for their effects on reproduction at very low doses. Results show that effects recorded varied, sometimes showing a more pronounced impact on the reproduction (10 out of 40 cases) than on the vegetative parts, but not always.

Effects on plants have seldom been assessed on seedling growth of the subsequent generation as in case 4 above (but see Blackburn and Boutin, 2003).

Strandberg et al. (2012) studied the importance of selection of the appropriate endpoint relative to time of exposure using four non-target plants including two annual (*Silene noctiflora*, *Geum molle*) and two taxonomically related perennial species (*S. vulgaris*, *G. robertianum*) with three herbicides (glyphosate, metsulfuron methyl and mecoprop-P) exposed at both vegetative (four- to six-leaf stage) and reproductive stages. The calculated EC_{50} revealed that, regardless of the phenological stage during exposure, seed production was the most sensitive endpoint. Furthermore, biomass was found to not be a useful endpoint for these species when plants were exposed at the reproductive stage. At this stage, plants did not grow any longer and often even exhibited lower total biomass.

Using reproductive endpoints such as seed production or seed quality and germinability (preferably both) requires special attention on plant pollination. Hand pollination may be carried out but this may not be ideal in every case. Most herbicides have the potential, through their mode of action, to influence plant metabolites. It is not known for how long such changes in plant metabolites may be present during the recovery process. Glyphosate, for example, blocks the synthesis of three essential aromatic amino acids: phenylalanine, tyrosine and tryptophan. However, the content of phenylalanine in pollen and nectar has been shown to be an important factor that seems to determine bees' behaviour when searching for food (Petanidou et al., 2006). Plants that have been exposed to glyphosate may have reduced content of phenylalanine which may render the flowers less attractive to pollinators. Therefore, manual pollination may lead to overestimation of the reproductive output. However, this can be resolved by using self-pollinated species where possible.

In many European countries, herbicides are sprayed several times during the growing season, whereas, in other countries, for instance in North America, herbicides are usually sprayed once at the onset of the growing season. Two studies were available in the literature that established the phenological stage of plant species present in boundaries at the time of herbicide spray. In 2007 and 2008, 40 hedgerows were surveyed in two studies conducted in Denmark in organic and conventional farming systems (Strandberg et al., 2012; Boutin et al., 2014). All hedgerows were at least 400 m in length and were

located in East-Jutland, Denmark. Sampling of phenological stages of all vascular plants was carried out monthly in 15 quadrats per hedgerow, from early May to mid-September. Overall, the hedgerows surveyed comprised 193 species. There were plants in flower from May (10–13 %) until September (24–32 %). The greatest flowering occurred in July, with around 60 % of species in flower. Many species were in flower for several months. All species but one flowered at some point during the growing season. Better represented families with the most species in flower were the Asteraceae (33 species), the Poaceae (30), the Fabaceae (12) and the Caryophyllaceae (11). Herbicide applications in Denmark takes place from March to mid-June and then in September and October, which does not coincide with the peak flowering time. Nevertheless, up to 40 % of species were in flower during normal spray application in May, June and September.

A study performed in eastern Canada in woodlots was used to document plant species composition, phenological stages prior to spray and herbicide effects after spray (Boutin et al., 2014). The study was conducted during three years in three woodlots in Southwestern Ontario in Canada (42.58N, 81.14W). The three woodlots were abutted to fields planted with soybean (*Glycine* max) in 1993, maize (*Zea mays*) in 1994 and wheat (*Triticum sativum*) and maize in 1996. Herbicides (imazethapyr in 1993, dicamba in 1994 and MCPA in 1996) were sprayed under normal operational conditions by the field owner in May of each year. Plants in woodlots were surveyed before (May: usually one day before spray or a few days prior to spray depending on the weather) and after (May, June and July) herbicide application. All vegetation below 2 m in height was surveyed for community composition, phenological stage (vegetative or flowering) and symptoms of herbicidal impact (comparing qualitative visual assessment prior and after the spray event) in six quadrats (six distances from field edges) along 10 transects perpendicular to the fields.

A total of 104 species were identified in the three woodlots during the three years of the study. Of these, 34 species from 20 families were found flowering prior to spray operations, including several species of forbs, graminoids and shrubs. These 34 species represents 33 % of the total number of species inventoried. Families with the most species in flower were the Asteraceae, Rosaceae and Violaceae. Between 12.5 and 28.5 % of the species were in flower per quadrat at the time of spray with no consistent trends between distances.

In the same study, it was shown that several species were affected by herbicide spray. A total of 34 species were found with marked herbicide effects including epinasty; leaf mottling, withering and yellowing; leaf and stem bending and twisting; necrosis; and bud malformations. Of these, 13 species were in flower during herbicide application and an additional five tree species were probably in flower, since they are early spring bloomers and exhibited symptoms consistent with pesticide poisoning.

It should be stressed that the choice of endpoints should strongly depend on the growth stage of plants at exposure. It was clearly demonstrated that plant exposure to herbicide regularly occurs at various phenological stages, be it seedling, juvenile, vegetative or reproductive stages, and that sensitivity varies greatly depending on the phenological stage at exposure and whether it is species dependant.

Overall, tables shown in Appendices A and B demonstrate the importance of considering the reproductive endpoint in regulatory assessment.

4.3.1. Duration of the test

In current standard guidelines, the effects of PPPs applied in early phenological stages are recorded 14–28 days after application when most plant species are still at the vegetative stage at harvest (OECD, 2006a, b; US EPA, 2012a,b,c): a rather short time frame. Plant sensitivity to herbicides depends on the phenological stage at spray application but also on the phenological stage used as the effect endpoint. However, overall effect on the whole life cycle is the crucial factor for the persistence of a natural plant population. This includes germinating seeds, seedling and juvenile stages, flowering, seed production and the germination rates of these seeds (F1).

Rokich et al. (2009) used a succession of five experimental settings to get a throughout picture of the effects of the herbicide fluazifop-p-butyl (848.0 and 844.8 L a.i./ha), a grass selective herbicide used for the control of introduced grasses. Plants of different developmental stages were exposed and subsequent effects on seed germination, seedling emergence, growth and health were recorded. In total, 13 different species were investigated, with monocotyledon and dicotyledon species representing native species as well as ones introduced to southwest Australia. Five different experiments were conducted: (1) application on seeds on filter paper and monitoring of seed germination, radical length and health for four weeks; (2) application on seeds buried 10 and 20 mm deep in soil and monitoring seedling emergence and seedling height for eight weeks; (3) application on seeds buried 10 mm deep in sand and monitoring of seedling emergence and seedling height for nine weeks (six monitoring weeks); (4) foliar and soil application on potted plants (3–4 months old) and monitoring of plant height and plant health for four weeks; and (5) application on potted plants (4–5 months old) and monitoring of plant height and plant health for six weeks. Studies showed that spray events in early stages could result in effects on flowering and seed production at even lower rates than those causing effects on mortality and juvenile growth. This is not evaluated in standard toxicity testing.

Additionally, there are several field studies investigating the effects of herbicide drift on yield or reproductive responses of plants, some of them including effects on germinability of F1-generation seeds. An overview is provided by the reviews of Obrigawitch et al. (1998) and Olszyk et al. (2004). In the reviewed studies, mainly crop species were treated with different herbicide rates and over two-thirds of these studies indicated reproductive or development effects (yield reduction) at less than field application rates. However, since the studies dealt with the negative effects of herbicide drift on neighbouring crops and the related yield loss, there is only a limited number of plant species, mainly crops, tested in these studies.

Some studies evaluated the effect of PPPs on the subsequent germinability of F1 seeds and seedling growth of non-crop plants. Blackburn and Boutin (2003) and literature reviewed therein showed that germination of seeds produced by plants exposed to the broad-spectrum herbicide glyphosate was affected and that this was dependent on the degree of seed maturity at exposure and on the level of herbicide used. Early reproductive growth stage and seed moisture level seem to be key parameters. The literature reviewed showed effects of 11 plant species of four families on seed germination and seed development and that effects depend on glyphosate application rate and timing of application. The authors tested 11 species of different families at 0.1, 1, 10 and 100 % of the recommended application rate (890 g a.i./ha) sprayed near seed maturity. The rates do account for rates reaching off-crop plants via spray drift. Of the species tested, 7 of the 11 showed significant effects of glyphosate treatment on germination of F1 generation and/or seedling growth. Overall, the data presented in the study (from literature and experimental components) showed that, for 19 species of seven families sprayed with glyphosate, there is an effect on vegetative parts as well as on the next generation.

Rinella et al. (2010b) investigated the possibility of using plant growth regulators to control invasive annual grasses by depleting their short-living seed banks. It was reported that reduced cereal grain yield was observed when plant growth regulators were applied in cereal (*Triticum aestivum*, *Hordeum vulgare*, *Zea mays* and *Avena sativa*) fields at late growth stages (Friesen et al., 1968; Sikkema et al., 2007; Rinella et al., 2010b). In two greenhouse studies, Rinella et al (2010a) investigated the effect of 2,4-D (at 1.12 kg a.i./ha), dicamba (at 0.56 kg a.i./ha) and picloram (at 0.42 kg a.i./ha) on the invasive annual grass *Bromus japonicus* at four different developmental stages (seedling, initiation of internode elongation, boot, and heading; Rinella et al., 2010a). The rates chosen were those commonly used for broadleaf weed control. Effects on seed production were consistently more pronounced than the effects on biomass. Picloram caused effects on biomass only if applied at the seedling and internode stage; dicamba showed negative effects on biomass only if applied on the seedling stage. In contrast, the application of picloram and dicamba at all stages tested resulted in reduced seed production. Picloram applied at the internode stage or later caused total seed loss.

In a field study conducted in eastern Montana, USA, Rinella et al. (2010b) applied picloram (0.42 kg a.i./ha) and aminopyralid (0.07 and 0.12 kg a.i./ha) on *Bromus japonicus* at three different growth

stages (initiation of internode elongation, boot, and shortly after heading; Rinella et al., 2010b). Less than 5 % germinability could be found for all combinations of timing (growth stage) and rate.

In the light of the limited results presented here and in section 6, it may be concluded that the whole life cycle has to be considered to properly assess impacts of herbicides on populations in agricultural areas. Herbicide as well as plant growth regulator application may affect later phenological stages and the juvenile stage is not always the most sensitive one to test. Further research is needed to strengthen these important findings.

4.4. Conclusion and highlights

- Two OECD guidelines (2006a, b) are available for testing pesticides on plants under greenhouse- or growth chamber-controlled conditions; however, these tests are of limited duration and do not allow for the assessment of the most sensitive endpoints such as seed production and germinability.
- The US EPA provide four documents (2012a, b, c, e) which describe testing of pesticides under controlled and field conditions.
- The use of the plant screening data (efficacy and crop margin of safety data) can be used in pesticide risk assessment, even though testing is not conducted under GLP.
- Plant screening data provide valuable information and easily accessible data on a wide range of species on effects of herbicides (and other pesticides) on non-target plants that are constituents of wildlife habitats.
- Plants sprayed at the juvenile stage as in routine regulatory testing may show herbicide injury in their vegetative parts and/or in their reproductive output.
- Plant sensitivity to herbicides varies depending on their phenological stage at the time of spray.
- Four cases can occur when herbicides are sprayed in crop fields and misplacement occurs into boundaries: non-target plants are at the juvenile stage and effects occur on (1) the vegetative parts or (2) later on the reproductive parts; moreover, herbicides reach plants when they are at the reproductive stage and effects occur on (3) the reproduction or (4) later on the next generation (seedling growth).
- Studies in Europe and North America have revealed that non-target plant species in boundaries may be at both the vegetative stage and the reproductive stage during spray events.
- ER_{50}s calculated for the vegetative stage (short-term toxicity), and the ER_{10} for the reproductive stage (chronic toxicity), were used to compute extrapolation factors for plants sprayed at the juvenile stage.
- The available data on both vegetative and reproductive (seed production) endpoints (see Appendix A) demonstrate the importance of considering a reproductive endpoint in regulatory assessment or use of an extrapolation factor to compensate for higher sensitivity of reproductive endpoints.
- Plants sprayed at the reproductive stage exhibited significant effects on both the vegetative and reproductive parts, and, in 10 out of 40 cases, effects were more pronounced on the reproduction than on the vegetative parts.
- It is important to carefully consider the choice of endpoints. It seems apparent that vegetative biomass is not always a useful endpoint when dealing with effects on later stages, e.g. the reproduction stage and exposure at the flowering stage; here seed production is more suitable.
- When studying reproductive endpoints it is important to ensure that pollination takes place either by doing the pollination manually or by making the required vector/animal for successful pollination available.

5. Exposure: off-field emission/deposition routes

5.1. Linking exposure to effects

EFSA (2010) indicated that the first step for linking exposure to effects is to define the type of ecotoxicologically relevant exposure concentrations (ERCs) that are needed. When linking exposure to effects for risk assessment, the same ERC should be used for both field exposure estimates and effect estimates. However, the relevant test endpoints for non-target plants are traditionally expressed in mass per area rather than concentrations. Consequently, for NTTPs, field exposure has to be expressed in the same unit (mass per area). Therefore, all methodologies recommended in section 5 are given as loads in mass per area rather than concentrations.

The Vegetative Vigour Test (e.g. OECD 227) for NTTPs considers mainly uptake via the aboveground plant parts because plants are oversprayed. In this test, exposure via roots will also occur, but to a lesser extent. The exposure in the Seedling Emergence Test (e.g. OECD 208) is via the radicles and, as seedlings grow, is via roots and emerging cotyledons and leaves; hence, the concentration in soil is relevant for this test.

In both the Vegetative Vigour and Seedling Emergence tests, plants or soils are oversprayed to mimic a spray event, and the endpoints are expressed in applied rates per hectare. The Seedling Emergence Test can also be conducted with spiked soil, and the endpoints expressed as concentrations in the soil, e.g. for special application types as incorporation in soil (these concentrations can be recalculated to applied rates per hectare). Consequently, the calculation of field exposure in section 5 focuses on droplet, vapour and dust drift.

Since the exposure in the Vegetative Vigour Test is expressed as an oversprayed dose (without considering any later dissipation on the plant surface), it is recommended that the maximum (initial) load be considered when calculating the respective field exposure rather than time-averaged figures. As a consequence of the previous recommendation, the relevant multiple application factor (necessary when the pesticide is applied more than once within a season) should not be based on residue adding but on effect adding as explained in the respective appendix E. If the compound is persistent in soil and exposure is mainly via the roots, then a residue-based multiple application factor (MAF) is more applicable. This approach will be elaborated on in a later guidance document.

In addition, if it is known that the pesticide is taken up by the plant root system, run-off should be considered as an additional exposure route using the methodology explained in section 5.6 and the results should be compared with seedling emergence concentration/response tests.

5.2. Exposure routes

The presence of PPPs on off-field non-target surfaces (plants, arthropods, bees, etc.) is a combination of three processes during and after the application of the compounds in the field: (1) the emission of the applied product out of the field edges by drift and runoff, (2) the deposition of the emitted amounts onto the off-field surfaces, and (3) dissipation processes from the non-target surface. However, when assessing the exposure of non-target plants, the dissipation processes from the non-target surface do not need to be considered if the mode of entry is through the plant leaf system. In current EU and national risk assessment, drift is considered to be the most important factor for off-field emissions to non-target surfaces. Losses owing to surface run-off may contribute to the contamination of non-target terrestrial ecosystems in the neighbourhood of agricultural areas if the mode of entry is uptake by the plant root system. Other emission routes such as leaching and drainage are generally not considered as direct emission routes. Drift is defined as droplet drift, but vapour drift and dust drift are also considered to be important emissions in some particular cases. “Deposition” on non-target surfaces is defined as the entry path for transport of airborne or waterborne substances from the air to the non-target surface, i.e. to an aquatic or terrestrial compartment or to non-target plants, arthropods, bees, etc. Dry and wet deposition should be considered separately because they are subject to different atmospheric and physical processes. Surface run-off may contribute to the contamination of non-target

terrestrial ecosystems in the neighbourhood of agricultural areas and is relevant for non-target plants if the compound is taken up by the plant root system. The soil concentrations should be based on the most recent scenarios published by EFSA (2012). The pore water scenarios should be considered as the preferred option, since soil water concentration is the key parameter for uptake via the plant root system.

In some cases, exposure of non-target surfaces is considered negligible and is not further assessed, e.g. in the case of rodenticides and substances used for wound protection. If substances are applied to stored products or in greenhouses, deposition caused by droplet drift, dust drift or run-off can be considered negligible too. However, for volatile compounds, vapour drift may nevertheless be significant and could be assessed in a similar way as in the field.

As mentioned in EFSA (2010), the exposure estimate should preferably apply to a given percentile of the concentration distribution (usually the 90th percentile) of the treated fields. Developing an exposure scenario for a given percentile requires simulating the concentration distribution in the entire target area (e.g. EFSA, 2012). The model for simulating this concentration distribution should preferably include all relevant exposure routes (i.e. spray drift deposition, vapour drift deposition, dust drift deposition and surface run-off). Since such models are not yet available for regulatory purposes at the European level, the simplifying assumption is made that the individual exposure routes can be assessed separately.

5.3. Droplet drift/deposition

Spray drift is defined as the part of the applied product that leaves the treated field through the air because of air currents during the application of the PPP. These droplet drift emissions do not include emissions by volatilisation. Droplet drift is considered to be a short-distance process (0–30 m) and occurs only during and shortly after application (i.e. within a few minutes actually defined by the time between spraying and collection of samples during drift experiments).

Droplet drift is not compound specific but is mainly dependent on droplet size, wind speed, wind direction and crop and spray-boom height during spraying. The spray drift is calculated on the basis of spray drift tables, which give the deposition as a percentage of pesticide application rate deposited at a given distance from the last crop row as a function of crop type (arable crops, fruits, grapes, hops and vegetables), crop stage (early or late) and spraying technique. Different spray drift curves are available (Southcombe et al., 1997; Rautmann et al., 2001; van de Zande et al., 2012, 2014). These spray drift curves were obtained from deposit measurements on artificial receptors (e.g. filter paper strips) on soil level. Most off-field emissions are calculated for deposition on surface water or soil. However, interception by non-target plants can be influenced in a different way because droplets have less contact with leaf surface owing to lower velocity and because of the presence of a laminar air layer on the leaf surface which influences contact. Moreover, the height and structure of the canopy is different from that of bare soil. For example, Kjær et al. (2014) demonstrated that spray drift deposition decreased with height in the plant canopy and that the effect of height is different at larger distances from the field. The PPR Panel did not review datasets to quantify these effects and assumes that the current methodology to assess spray drift deposition (FOCUS, 2001) will continue to be used for exposure assessment at EU level until better alternatives become available.

Currently, estimation of spray drift deposition is based on the values given by Rautmann (2001). These values apply to 90th percentile conditions. However, in a workshop on harmonisation of European drift values (Huijsmans and van de Zande, 2011; van de Zande et al., 2014), it was concluded that spray drift values in field crops originating from recent research were considerably higher than the values used by FOCUS (2001). These differences were particularly important at short distances (0–3 m) of the treated crop (Table 4) and were caused by differences in the selection of datasets to fit the reference deposition curves upon. Crop and spray-boom height during application of the pesticide are other important reasons for differences in spray deposition. For this reason, van de

Zande et al. (2012, 2014) suggest using a different spray drift curve for developed crops than for short crops.

Table 4: Estimated spray drift deposition for field crops (% of in-field target deposition) downwind of a sprayed (downwards) bare soil surface and a crop situation based on joined spray drift data from Germany and the Netherlands (van de Zande et al., 2014) and FOCUS (2001)

	Distance from the treated area of the crop (m)			
	1	**3**	**5**	**10**
van de Zande et al. (2014): crop	42.5	6.7	2.8	0.88
van de Zande et al. (2014): bare soil	7.77	1.93	1.01	0.42
FOCUS (2001): crop and bare soil	2.77	0.95	0.57	0.29

As stated in section 2.3.2, the risk assessment for the off-field area could consist of two steps. In the first step, the exposure would be based on the in-field risk assessment, i.e. the drift deposition would be 100 % of the sprayed dose rate. If the protection goal for the off-field area would not be met in this step, risk mitigation options would have to be assessed in a follow-up step. Options to mitigate spray drift deposition to off-field areas include (1) the use of spray drift reducing techniques, and (2) the establishment of non-spray buffer strips, with or without crop. Since spray drift deposition decreases with both distance and drift-reducing technique class, spray drift mitigation options could be evaluated using a matrix. Spray drift deposition could, for example, be evaluated first for the standard spraying technique, second for drift-reducing techniques and measures and third for all spray techniques with stepwise wider buffer strip.

Spray drift deposition differs between crop types (grass and bare soil, field crops, fruit crops, vines and hops) and crop development stage. For this reason, a spray drift deposition curve and hence an evaluation matrix is needed for each combination of crop type and crop development stage, or classes of these. For estimating spray drift deposition onto surface waters, spray drift deposition curves were developed by the FOCUS Surface Water working group (FOCUS, 2001) for many major crops. Harmonised European drift curves are currently only available for bare soil, grass and fully developed arable field crops (van de Zande et al., 2014, draft version); spray drift curves for fruit crops are expected to become available by autumn 2014. However, for vine and hops, no updated values are foreseen in the near future. In this situation, the PPR Panel recommends the evaluation of new spray drift curves when they become available and to start revising the spray drift assessment methodology accordingly. For the time being, the PPR panel recommends the use of the current assessment based on FOCUS (2001). However, please note that the exposure assessment for all environmental compartments in which spray drift is relevant would benefit from this revision. A summary of the current spray drift assessment methodology and guidance for how to calculate the maximum exposure rate is given in Appendix D.

5.4. Vapour drift

Vapour drift can occur by (1) evaporation of the solvent from small spray droplets which are still present as “drift” after application and (2) post-application evaporation of the spray deposits from treated plant/soil surfaces. Vapour drift deposition is usually short to medium range (0–1 000 m). Most emission by volatilisation occurs during the first 24 hours after application caused by spray droplet evaporation. However, this process may continue for several days or weeks after treatment (Bedos et al., 2002). Evaporation of the leaf/soil deposits is dependent on the active ingredient properties, such as volatility, and interaction with leaves. Volatilisation from plant surfaces is one of the main pathways of pesticide emission to the environment and normally is greater than volatilisation from soils because plants have fewer sorption sites than soil.

The main factors controlling pesticide volatilisation are the physicochemical properties of the pesticide (in particular vapour pressure), agricultural practices (time and type of application), soil or plant physical properties and meteorological conditions (during and after application). Several models for

vapour drift emissions were evaluated by the FOCUS Air working group (FOCUS, 2008). They concluded that none of the models available completely fulfilled the requirements for use within a regulatory context. For pragmatic reasons, FOCUS (2008) recommended using the EVA 2 model for calculating the deposition after volatilisation for short-range transport. Subsequently, the PPR Panel evaluated this model and came to the conclusion that the recommendations regarding the use of the EVA 2 model are scientifically not robust enough (EFSA, 2007). The PPR Panel further came to the conclusion that the recommended model does not give realistic worst-case exposure estimates. Therefore, the PPR Panel recommends improving the estimation of vapour drift deposition by the EVA 2 model and also investigating the option to use alternative modelling approaches, since these have now become available.

One modelling approach that has potential for assessing vapour drift exposure of non-target plants has been developed to improve the assessment of vapour exposure of workers, residents and bystanders (van den Berg et al., 2013; see also www.browseproject.eu). They used the PEARL model (Tiktak et al. 2000) to calculate the emission from the crop into the air and the OPS model (Jaarsveld, 2004) to calculate the subsequent transport via the air based on real meteorological data for a five-year period. Based on this, they developed 90th percentile (realistic worst-case) exposure scenarios based on the temperature distribution in Europe, since vapour pressure increases with temperature.

FOCUS (2008) stated that volatilisation is only relevant for compounds with a vapour pressure higher than 10^{-4} Pa when applied to the soil and for compounds with a vapour pressure higher than 10^{-5} Pa when applied to the crop. In this context, it is worth noting that, whereas it is possible to minimise droplet drift emission to the off-crop area using appropriate application techniques (e.g. drift-reducing nozzles, buffer zones), this does not apply to volatilisation, since this process is mainly driven by pesticide and crop properties. The relative contribution of vapour drift deposition is demonstrated in Tables 7 and 8. These tables summarise calculations with the EVA 2 model for a compound with medium volatility (vapour pressure of 5×10^{-3} Pa) assuming various crop interception fractions in the field. The consideration of crop interception is necessary, since volatilisation from the crop canopy is on average three times higher than volatilisation from the soil surface. For the crop interception values, the most recent number are recommended (EFSA, 2014).

Table 5: Droplet drift and volatilisation deposits for arable field crops calculated with EVA 2 [(a)]

Distance (m) [(b)]	**Droplet drift (µg/m²)**	**Cumulative volatilisation deposits over 24 h (µg/m²) dependent on crop interception in field**			
		No interception	**25 % interception**	**70 % interception**	**90 % interception**
1	2 770	518	778	1 244	1 451
3	943	465	697	1 116	1 302
5	570	417	625	1 000	1 167
10	290	318	476	762	889
15	200	242	363	580	677
20	150	184	276	442	516
30	100	107	160	256	299
50	60	36	54	86	101
100	30	2	4	6	7

(a): Application dose 1 kg/ha, vapour pressure 5×10^{-3} Pa.
(b): From last row of treated crop.

Table 6: Droplet drift and volatilisation deposits in orchards (early) calculated with EVA 2 [(a)]

Distance (m) [(b)]	Droplet drift (µg/m²)	Cumulative volatilisation deposits over 24 h (µg/m²) dependent on crop interception in field			
		No interception	25 % interception	70 % interception	90 % interception
1	–	1 814	534	548	670
3	29 200	1 627	479	492	601
5	19 890	1 459	429	441	539
10	11 810	1 111	327	336	410
15	5 550	846	249	256	313
20	2 770	645	190	195	238
30	1 040	374	110	113	138
50	300	126	37	38	46
100	60	8	2	2	3

(a): Application dose 1 kg/ha, vapour pressure 5×10^{-3} Pa.
(b): From last row of treated crop.

It should be noted that the depositions of vapour drift in the Tables 5 and 6 are the cumulative exposure over 24 hours. They are compared with the deposition of spray drift which can be considered as an instantaneous event actually defined by the time between spraying and collection of samples during drift experiments (usually 15 minutes).

The relationship between vapour drift deposition and distance is described in EVA 2 by an exponential function (equation 5.4.1):

$$D_V = D_{V1} * \exp[-0.05446 * (d - 1)] \qquad (5.4.1)$$

D_V: relative volatilisation deposit (% of application dose)

d: distance from the edge of field (m)

D_{V1}: relative volatilisation deposit at 1 m distance from the edge of field (% of application dose)

The relative volatilisation deposit is given for three classes (Table 9). EVA 2 assumes volatilisation deposits at 1 m being dependent on the vapour pressure at 20 °C. Table 7 shows the respective deposits at 1 m of edge of field from plant and soil surfaces.

Table 7: Relative volatilisation deposits D_{V1} at 1 m distance used by EVA 2

Vapour pressure (vp) range at 20 °C	Relative volatilisation from canopy, deposit at 1 m (% of application dose)	Relative volatilisation from soil surface, deposit at 1 m (% of application dose)
$vp < 10^{-5}$ Pa (plant) $vp < 10^{-4}$ Pa (soil)	0.00	0.00
10^{-4} Pa > vp > 10^{-5} Pa	0.09	0.03
$5x10^{-3}$ Pa > vp ≥ 10^{-4} Pa	0.22	0.073
$5x10^{-3}$ Pa > vp ≥ 10^{-2} Pa	1.56	0.52
$vp > 10^{-2}$	Case by case	Case by case

As shown in Table 6, the procedure does not hold for pesticides with a very high vapour pressure (> 10^{-2} Pa at 20 °C, see section 5.4.4 in FOCUS (2008)). In these specific cases, risk managers might want to request dedicated field experiments.

If the deposition rate exceeds the maximum acceptable dose, risk assessors might want to establish a buffer strip. Similar to the calculation of drift deposits, equation 5.4.1 does not directly allow for the calculation of buffer strip width necessary to meet maximum acceptable deposits. However, this is possible when using equation 5.4.2, which can be obtained from the equation 5.4.1 after transformation.

$$d = 1 + \frac{\ln\left(\frac{3 \cdot 100 \cdot MaxLoad}{AppDose \cdot D_{VP1} \cdot \left(2 \cdot \frac{I_C}{100} + 1\right)}\right)}{0.05446\ m^{-1}} \qquad (5.4.2)$$

d: necessary width of buffer strip (m)

MaxLoad: maximum acceptable load (kg/ha)

AppDose: application dose (kg/ha)

D_{VP1}: relative volatilisation deposit from canopy at 1 m distance from the edge of field (% of application dose)

I_C: crop interception (field crop) during application (%)

Table 8 shows examples for different volatility classes when 1 kg/ha was sprayed and the maximum acceptable load was calculated to be 0.01 kg/ha. Please note that this table is just an example; if, for example, the maximum acceptable load is lower, then the buffer zones would be larger as well.

Table 8: Example of necessary buffers (m) to prevent non-acceptable volatilisation deposits [(a)]

Vapour pressure (vp) range at 20 °C	Relative volatilisation from canopy, deposit at 1 m (% of application dose)	Necessary distance (m) no interception (in field crop)	Necessary distance (m) 100 % interception (in field crop)
vp < 10^{-5} Pa (plant) vp < 10^{-4} Pa (soil)	0.00	No buffer	No buffer
10^{-4} Pa > vp > 10^{-5} Pa	0.09	No buffer	No buffer
$5x10^{-3}$ Pa > vp ≥ 10^{-4} Pa	0.22	No buffer	No buffer
vp ≥ $5x10^{-3}$ Pa	1.56	No buffer	9.17
vp > 10^{-2}	Case by case	Case by case	Case by case

(a): Application dose 1 kg/ha, maximum acceptable load 0.01 kg/ha.

In Table 8, a buffer zone was calculated only for the compounds having vapour pressures above $5x10^{-3}$ Pa and when the application was targeted fully at the (target) canopy. This demonstrates that, in most situations, deposition caused by droplet drift will be the dominant entry route rather than volatilisation deposits.

5.5. Particulate drift

Particulate drift can occur due to (1) the application of dust from dustable powder formulations (e.g. sulphur dusting in vineyards), (2) dust formation during non-spray applications (NSA), e.g. granules (fertiliser–herbicide combinations for application in lawns), treated seeds, or (3) soil dust with adsorbed pesticide deposits. However, the latter emission is not considered to be a direct emission route.

Particulate drift happens generally over a short range and in short periods after application and is thus comparable to droplet drift. The main driving force is the particle size/weight of the dust particles.

The EFSA opinion on NSAs (EFSA, 2004) gives guidance for the exposure assessment of NSAs. The main conclusions and recommendations for dust drift are the following:

1. Dust in NSAs is a relevant route of exposure for surface water.
2. Drift is a non-relevant route for granules and seed treatment applied in furrow and buried immediately, as well as for coated seeds.
3. However, broadcast granular applications, even with subsequent incorporation, can form dust drift, which can have comparable effects as spray drift.
4. Broadcast application of treated seeds (with and without subsequent incorporation) is also considered to be a relevant route of exposure.

According to the EFSA opinion on NSAs (EFSA, 2004), dust drift can be handled by FOCUSsw models with adjustment of the normal default inputs in such a way that an evaluated dust drift value is entered. For default values, adapted spray drift models can be used to estimate dry deposition from dust by taking into account a number of specific underlying criteria.

With regard to seed treatment, the European Commission recently prepared a document which includes experimental data from dust drift deposition for different crops (EC, 2014). It is stated that, for NTTPs, no direct exposure via the soil as a result of treated seeds is expected. Furthermore, it is not expected that a risk assessment for non-target plants via dust is considered necessary, since herbicides are not used for seed treatment. According to the document the only exception should be that screening data indicates that the product may have adverse effects on plants. If that is the case, there is however a problem when considering the experimental studies on dust drift deposition in this document. In contrast to the standard FOCUS values, the evaluation performed in EC (2014) does not give dependencies between the deposition and the application dose, which would be necessary to define safe distances dependent on the application rate.

5.6. Run-off entries

The assessment of pesticide movement to surface water caused by run-off is currently a key process in European risk assessment. The recommended methodology as described by FOCUS (2001) follows a tiered approach. Run-off occurs after heavy rainfall events which may transport residues of the active substance or transformation products either dissolved in the water or sorbed to the eroded sediment phase to the non-target area.

For the estimation of run-off and erosion losses leaving the edge of field, several models are available, e.g. the models used in the different tiers of FOCUS surface water (FOCUS, 2001). At tier II, pesticide losses by run-off as summarised in Table 9 are considered.

Table 9: Step 2: pesticide losses by run-off and soil erosion according to FOCUS step 2

Region/season	% of soil residue
North/centre [(a)] Europe, October–February	5
North/centre [(a)] Europe, March–May	2
North/centre [(a)] Europe, June–September	2
South Europe, October–February	4
South Europe, March–May	4
South Europe, June–September	3

(a): According to FOCUS (2001) the number also reflects the situation in Northern France.

For pragmatic reasons, the losses due to run-off at step 2 were defined by FOCUS independently on sorption properties of the compound. According to FOCUS, they have been calibrated against the results of tier-III calculations. The key model for the estimation of run-off in FOCUS at tier III is PRZM. Reichenberger et al. (2007) made a probabilistic analysis of losses caused by run-off and erosion using the Pesticide Root Zone Model (PRZM) and analysed losses dependent on sorption. For run-off, the maximum losses were found for compounds with K_{OC} values in the range of 100–200 L/kg. For losses by soil erosion, the maximum numbers were found for compounds with maximum K_{OC} values. The results were evaluated by the German federal environmental protection agency and, meanwhile, were also implemented into their model EXPOSIT 3.0 used in German risk assessment for estimating pesticide losses caused by run-off (Umweltbundesamt, 2011). However, presently, this analysis is of use only in the central European zone, since only German environmental conditions were considered. However, it is recommended that the dependencies between important pesticide properties and run-off losses for all European zones be analysed in order to improve the information given in Table 3 by FOCUS (2001).

As mitigation measures for run-off entries reaching surface waters, EFSA PPR Panel (2013) recommends the use of vegetated buffer strips taken from FOCUS (2007) as summarised in Table 10.

Table 10: 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, 2007)

Buffer width (m)	Reduction of run-off (%)	Reduction of erosion (%)
5	37 [(a)]	55 [(a)]
10	60	80
15	72	90 [(a)]
20	80	95

(a): Suggested numbers for buffer strips up to 5 m and 15 m from the field.

These mitigation measures for surface water are directly connected with deposition of residues to the respective terrestrial ecosystems (vegetated buffer strips). Table 10 shows the total reduction at a certain distance from the field. In Table 11, the reduction factors are recalculated to give the reduction that is occurring at the respective distance only.

Table 11: Deposited fraction dependent on the position of the buffer strip (calculated from Table 5)

Buffer width (m)	Run-off fraction deposited (%)	Erosion fraction deposited (%)
0–5	37	55
5–10	23	25
10–15	12	10
15–20	8	5

Only substances with very high sorption constants are transported via the sediment. Since uptake of such substances by the plant root system is negligible, it is not expected that deposition via the sediment will cause adverse effects on NTTPs. Therefore, sorption via the sediment phase is not considered.

In the following equations, the numbers in Table 11 were transformed into mathematical expressions. It was not possible to describe the deposition of pesticide in water by a single exponential function.

Therefore, the following sequential exponential functions could be used to derive deposited fractions for any distance for the water phase:

$D_R = D_0 * f_R * \exp(-0.09163 * d)$ (d ≤ 10 m) (5.6.1)

$D_R = 0.8 * D_0 * f_R * \exp(-0.06931 * d)$ (d> 10 m) (5.6.2)

$$f_R = \frac{1}{1 + \left(c_{sed} \cdot \frac{K_{OC} \cdot C_{org}}{100} \cdot \right)} \quad (5.6.3)$$

D_R: relative deposit owing to run-off at a given distance (% of application dose)

D_0: relative pesticide loss owing to run-off (%, see Table 8)

d: distance from the edge of field (m)

f_R: fraction of pesticide transported in the water phase (–)

K_{OC}: sorption constant related to organic carbon (L/kg)

C_{org}: organic carbon content in soil (%), e.g. 2 %

c_{sed}: sediment particles in run-off (kg/L), e.g. 0.01 kg/L

Similar to the calculation of drift deposits, equations 5.6.1 to 5.6.3 do not directly allow the calculation of buffer strips necessary to meet maximum acceptable deposits. However, this is possible when using equations 5.6.4 and 5.6.5, which can be obtained from the equations above after some transformation, and assuming that the maximum acceptable load is calculated based on the application dose and the necessary reduction to meet the maximum acceptable load:

$$d = \frac{\ln\left(\frac{100 \cdot MaxLoad}{AppDose \cdot D_0 \cdot f_R} \right)}{0.09163\ m^{-1}} \quad \text{(d ≤ 10 m)} \quad (5.6.4)$$

$$d = \frac{\ln\left(\frac{100 \cdot MaxLoad}{0.8 \cdot AppDose \cdot D_0 \cdot f_R} \right)}{0.06931\ m^{-1}} \quad \text{(d > 10 m)} \quad (5.6.5)$$

d: necessary width of buffer strip (m)

MaxLoad: maximum acceptable load (kg/ha)

AppDose: application dose (kg/ha)

D_0: relative run-off/erosion loss at edge of field dependent on season and region (% of application dose)

f_R: fraction of pesticide transported in the water phase (–)

Table 12 shows examples for the acceptable distance in different seasons when 1 kg/ha was applied and the maximum acceptable load was calculated to be 0.01 kg/ha.

Table 12: Example of necessary run-off buffers (m) in different regions/seasons [(a)]

Region/season	% of soil residue leaving the field [(b)]	K_{OC} (L/kg)	Necessary distance (m)	Calculated with equation
North Europe, October–February	5	100	19.7	5.8b
North Europe, March–May	2	100	7.3	5.8a
North Europe, June–September	2	100	7.3	5.8a
South Europe, October–February	4	100	16.5	5.8b
South Europe, March–May	4	100	16.5	5.8b
South Europe, June–September	3	100	12.3	5.8b
North Europe, October–February	5	5 000	10	5.8a
North Europe, March–May	2	5 000	No buffer	5.8a
North Europe, June–September	2	5 000	No buffer	5.8a
South Europe, October–February	4	5 000	7.6	5.8a
South Europe, March–May	4	5 000	7.6	5.8a
South Europe, June–September	3	5 000	4.4	5.8a

(a): Application dose 1 kg/ha, maximum acceptable load 0.01 kg/ha, organic carbon in soil 2 %, concentration of suspended particles in run-off 0.01 kg/L.
(b): Degradation in soil before run-off event not considered.

Dependent on the K_{OC} of the compound, the deposition in the non-target area is calculated via the water phase (low to moderately sorbing compounds, equations 5.6.4 and 5.6.5). The results presented in Table 12 should explain the influence of the key input parameters on the width of the buffer zone as expressed in equations 5.6.1, 5.6.2 and 5.6.3. However, the actual risk assessment could be also performed by defining different classes for the different input parameters and by developing a matrix giving necessary distances for each combination of them.

5.7. Conclusion and highlights

- Drift during application is currently considered to be the most important factor for off-field emissions to non-target surfaces. Drift is normally defined as droplet drift, but vapour drift can also contribute in particular cases. Exposure models to calculate loadings caused by droplet and vapour drift are presently available. However, spray drift values in field crops originating from recent research were considerably higher than those currently used in exposure assessments at the EU level. The PPR Panel therefore recommends reviewing this new research and to update the spray drift models after this review has been carried out.
- Dust drift is considered to be an important emission route in particular cases. However, no validated models are available so far. As dust drift of very small particles can behave in a similar way as vapour drift, it can be proposed as a starting point for dust deposition on soil (EFSA, 2012).
- Experiences from the exposure assessments of surface waters show that also surface run-off may contribute significantly to the contamination of non-target terrestrial ecosystems in the neighbourhood of agricultural areas. Models to estimate run-off losses are available and used for the assessment of the aquatic environment. However, the information on vegetated buffer strips used currently in the aquatic risk assessment has to be re-evaluated with regard to worst-

case situations for non-target plants. The exposure via soil residues is only of relevance for seedling emergence and for root uptake.

- In order to improve the linking of exposure to effects, more effort is required in predicting/measuring the actual load of PPPs that is reaching NTTPs (plant surfaces or internal concentrations).

6. Higher tier assessment based on refined laboratory, semi-field and field studies

6.1. Effect refinements

In this opinion, no proposal is made for a risk assessment scheme (this step will be done in a later phase when a new guidance document has been developed). Consequently, no prescribed steps for higher tiers are proposed. It is not clear whether in the future a stepwise approach will be provided for higher tier studies because each legislative problem does not necessarily have the same solution or will lead to the same fixed next step in the risk assessment scheme. In cases when the information provided for a first step risk assessment is not enough to predict safe use of the compound or formulation, it is advisable to look for the real problem for which an answer is needed (e.g. a potential reproductive problem or problem with false positives). Even before that, implementing a mitigation measure could be considered to prevent additional higher tier testing with all its additional uncertainties. To date, mainly the basic studies which are needed (required in the legislation) are performed for non-target plant risk assessment and the registration of herbicides. The result of this policy is that higher tier studies have been conducted only occasionally (Olszyk et al., 2004; Schmitz et al., 2013b) and no standard protocol for such studies is available. Higher tier assessment is not required if the predicted risk based on the basic studies could be managed by risk mitigation strategies. However, to refine the risk assessment, a higher tier study, including multi-species experiments in the greenhouse and/or field experiments, can be performed. Therefore, notifiers of pesticides might wish to discuss the study protocol and details on the test design with the responsible authority of a Rapporteur Member State.

Generally, at tier III, effects on non-target plants should be observed under more realistic conditions than for tier II studies. This, however, may include many different aspects of the study design and both biotic (e.g. species interactions) and abiotic (e.g. drift exposure, climate) test conditions may be more realistic. This section aims to summarise current knowledge on relevant studies for assessment of effects of herbicides on NTTPs at higher tiers with emphasis on study design and it includes:

1. additional laboratory tests including reproductive endpoints (section 6.1.1);
2. greenhouse tests including species interactions (section 6.1.2);
3. comparison of effects under greenhouse and outdoor/field conditions (section 6.1.3);
4. phytometer experiments where single species in pots or microcosms are placed in the field, i.e. under realistic spray drift conditions (section 6.1.4);
5. field experiments where one or several species (NTTPs) already growing in the field are exposed to herbicides at realistic doses (section 6.1.5);
6. field experiments where experimentally established species (NTTPs) growing in multi-species mixtures are exposed to herbicides at realistic doses (section 6.1.6).

Sections 2.9 and 2.10 describe the importance of long-term effect assessment, i.e. assessment of chronic effects for risk assessment for NTTPs, and also emphasises the importance of using reproductive endpoints for the assessment. Here we focus on available test protocols and study designs for this type of higher tier effect assessment in the laboratory/greenhouse.

The reviewed studies have mainly been carried out with herbicides, but it needs to be emphasised that the test designs might be used to study herbicidal effects of other PPPs.

6.1.1. Additional laboratory tests including reproductive endpoints

6.1.1.1. ISO 22030 (2005) for testing chronic toxicity to higher plants with provisions

The international standard ISO 22030 (2005) produced a protocol for "Soil quality—Biological methods—Chronic toxicity" in higher plants for evaluating the quality of contaminated soils. The evaluation is based on the assessment of two vascular plant species under controlled conditions. Two species with rapid life cycles are recommended: turnip rape (*Brassica rapa* CrGC syn. Rbr) and oat (*Avena sativa*). Both acute and chronic endpoints are measured (emergence, early growth and reproduction). Per pot, 10 seeds are sown, which is thinned to eight, and four plants are harvested at day 14 and at the end of the test (three to four weeks later for *B. rapa* and six weeks later for *A. sativa*). Other species with ecological or economic significance in certain regions of the world can be used from a list provided in ISO 11269-2 (2012). Reasons for selecting species other than oats and turnip rate have to be provided. Plants are watered via wicks. Recommended soils, according to OECD (1984) or ISO 11268-2 (2012), are suggested. The OECD soil is a sandy loam, loamy sand, sandy clay loam, or commercial potting or synthetic soil that contains up to 1.5 % organic carbon (approximately 3 % organic matter). The ISO soil is an artificial soil made of 10 % sphagnum peat, 20 % kaolinite clay and 70 % sand.

6.1.1.2. Other laboratory/greenhouse test designs for chronic assessment including reproductive endpoints

A number of greenhouse studies for assessment of the chronic effects of herbicides on NTTPs (or surrogate species) have been carried out, including studies that include assessment of the effects on reproduction.

Carpenter and Boutin (2010) used a greenhouse study of acute (short-term test) and chronic (long-term test) toxicity of sublethal concentrations of glufosinate ammonium on 10 crop species (four monocots from two families, six dicots from five families) and 10 wild species (four monocots from two families, six dicots from six families). The selected species included annuals/biennials (nine species) as well as perennial species (11 species) for both crops and wild species. Species selection followed the guidelines established by the US EPA (1996) and OECD (2006b). The test was carried out in the greenhouses at Environment Canada over a period of seven months (June to December). For the short-term tests, aboveground biomass was harvested 21 days after glufosinate exposure. In the long-term tests, plants were allowed to grow until fruit/seed production or until the controls showed signs of natural senescence or stress, at which point all plants of a given species were harvested. Seeds of all species were sown separately in trays of soil consisting of Promix-BX with Myccorrhize® and horticultural sand in a 3:1 mixture. All trays were placed in the greenhouse after sowing except for *Juncus dudleyi*, which was stratified in a 2–4 °C dark refrigerator before being placed in the greenhouse. One to two weeks after emergence, the seedlings were transplanted into pots (10 cm in diameter, 9 cm high). Plants were exposed to the herbicide when they had reached the three to six true leaf stage. Glufosinate ammonium was applied at eight doses following a geometric progression of 1.9, i.e. at 1, 1.9, 3.4, 6.9, 13, 24.8, 47 and 89 % of a 100 % dosage of 750 g a.i./ha. The applications were performed using a track spray booth equipped with a flat-fan nozzle. Six replicates, each consisting of one plant per pot, were used for each dose and control. This resulted in 108 plants per species (eight doses + controls × six replicates × two treatments (short vs. long term)). All plants were well hydrated prior to spraying in order to maintain humidity for glufosinate efficacy. Control plants were moved to the greenhouse, whereas newly sprayed plants were isolated from the main experimental greenhouses for a period of 24 hours to avoid potential volatilisation and drift of glufosinate. When returned to the greenhouse, all plants of a given species were randomised within set blocks and were rotated regularly. Plants for the long-term experiment were transplanted into larger pots (15 cm in diameter, 18 cm high) in order to minimise the risk of plants becoming pot bound. This

was done when the short-term plants had been harvested. At about 40 and 64 days after exposure, fertiliser was applied.

In 7 out of 12 cases where reproductive endpoints or a proxy were measured, reproductive endpoints were more sensitive than the short-term biomass endpoint. Tests lasted between 38 and 106 days for annuals that reproduced, whereas tests lasted between 66 to 139 days for perennials that reproduced. Even though several of the species required pollination to produce fruits/seeds, nothing is mentioned on pollination in the test description, since number of flowers was used as an endpoint in these cases.

A study using a similar protocol was conducted with chlorimuron ethyl (Carpenter et al., 2013). In this case, the experimental work lasted six months using nine terrestrial and eight wetland species. All the annual species flowered (n = 8), whereas six of the perennial species produced flowers or equivalent. Tests lasted between 7 and 23.5 weeks for annuals that reproduced, whereas it was between 12 and 23 weeks for perennials that reproduced. Of the 14 species tested, three exhibited more sensitivity on their reproduction then the short-term biomass.

Boutin et al. (2014) includes a Danish study of effects of fluroxypyr (Starane 180S) on flowering of two perennial species, *Taraxacum vulgare* and *Trifolium pratense*, whose flowers are very important to pollinating insects. Seeds of the two species were sown separately in trays of planting soil in the greenhouse and at the four- to six-leaf stage, four plants were transplanted to larger pots (15 cm in diameter, 18 cm high). To ensure that enough plants were available for the exposure experiment, 50 pots of each species were transplanted. When the plants were well established, about one and a half months later (mid-June), the pots were moved to outdoor conditions and stayed there until the plants were flowering; for dandelion, mid-April and, for red clover, mid-June the following year. Immediately before application of the herbicide, 16 pots of each species with the same number of flower buds and the same overall performance were selected and moved to the greenhouse. The plants were exposed to fluroxypyr at four doses: 0, 5, 25 and 100 % of the label rate of 144 g a.i./ha in a standard spraying chamber. Following application the plants remained in the greenhouse and stayed there to the end of the test, i.e. 60 days later when no more flowers bloomed and no more flower buds were formed. During this period, both the onset of flowering and the number of open flowers were recorded. All herbicide doses were sublethal to the plants but had visible effects even at the 5 % dose. The average cumulative number of flowers produced by *T. pratense* was severely impaired at all doses of fluroxypyr, whereas *T. vulgare* experienced effects at higher doses (> 5 % of label rate). Onset of flowering was also significantly delayed in both species except at the 5 % dose.

6.1.1.3. Specifications and recommendations for additional greenhouse (laboratory) tests for chronic effects

- The ISO methods are the only standardised test guidelines for higher tier tests and their usefulness for testing herbicide effects on NTTPs is very limited.
- Only two crop species with a very short life cycle are recommended for the ISO tests.
- These ISO methods call for testing either with contaminated soil or using a series of dilutions incorporated into control soil. There is no provision of testing overspray. The methodology described in other guidelines (OECD, 2006a,b, or US EPA, 2012a, b, c) could be used to accommodate this need.
- The artificial soil recommended for the ISO tests is a very poor soil in which plants do not grow well (10 % sphagnum peat, 20 % kaolin clay, 69 % sand). Sensitivity to toxicants in soil appeared to be more pronounced in poor soil (Kalsch et al., 2006; Smith et al., 2013). Recommendations in the more recent OECD guidelines (2006a,b) should be used.
- The test designs proposed by Carpenter and Boutin (2010) and Carpenter et al. (2013) represent greenhouse tests for long-term assessment of reproductive endpoints of both annual and perennial species.

- The test design used for the Danish study of plant flowering (Boutin et al., 2014) forms an important contribution to higher tier testing, as it is one of the first studies of herbicide effects on plant flowering. The study does not include annual plants. Potentially, annuals and perennials respond differently with respect to flowering and other reproductive endpoints when exposed to herbicides.
- If species selected for the tests, as is the case in all three studies including the ISO tests, require pollination to produce fertile seeds, pollination should be carried out either manually or by the introduction of relevant pollinators (honey bees, bumble bees and/or solitary bees).

6.1.2. Greenhouse experiments for assessment of species interactions

One way herbicides may affect plant community dynamics is by affecting the interactions between plant species, and both intra- and interspecific competition have been shown to be important for community structuring (Rees et al., 1996; Weiher et al., 1998; Silvertown et al., 1999; Gotelli and McCabe, 2002).

In highly fertile habitats where there are dense or high-biomass populations, competition between plants may be size asymmetric (Weiner, 1990). Typically, this occurs as a result of competition for light, which means that large individuals shade small ones, but not vice versa. When exposed to herbicides, small individuals might be shielded from herbicide exposure by larger plants (Riemens et al., 2008). At low plant densities, individual plants might recover from a low-concentration herbicide treatment. However, when growing together in dense mixtures, the competitive interactions among the species may increase the effect the herbicide treatment has on the growth of the most sensitive plant species, so that this species loses its position in the size/height hierarchy.

Effects of herbicides on plant interactions may be studied in two-species competition experiments or in microcosm experiments carried out in the greenhouse or outdoors. In such studies, which represent an important type of tier III assessment, two or more species (NTTPs or surrogate species) are grown together under controlled conditions and the herbicide is applied at doses covering 0–100 % of the field rate. Generally, these studies aim at comparing the sensitivity of species grown in monocultures (with intraspecific competition) with the sensitivity of the same species grown in competition with one or several other species.

6.1.2.1. Competition experiments in the greenhouse for assessment of species interactions

Dose–response experiments combined with response–surface competition experiments (Inouye, 2001), using a complete additive design as proposed by Cousens (1991), are well suited for studies of species interactions, although they are labour intensive and demanding with respect to greenhouse facilities. Damgaard et al. (2008) performed such an experiment with two annual weeds, *Capsella bursa-pastoris* and *Geranium dissectum*, exposed to three doses of mecoprop-P (0, 0.5 and 2.0 g a.i./ha). Seeds of both species were sown separately in trays and transplanted into boxes in geometric series of plant densities using a regular pattern. The density of each of the two species was 1, 4, 8, 16 or 32 plants per box, covering scenarios from no competition to high competition intensity. Each density of each species was combined with the four densities of the other species plus a high density with 64 plants per box, which was included to assess the intraspecific competition at high density. In total, 26 different plant compositions were used. The boxes were placed in the greenhouse and watered daily. The boxes were placed in three blocks with different densities, < 20, 20–40 and > 40 plants per box, to avoid neighbouring effects. Treatments were randomised within each block. Aboveground biomass was harvested three weeks after herbicide treatment. The plants in the outer 5 cm of the boxes were not included in the analyses.

In addition to the competition experiment, a standard dose–response experiment was carried out. The dose–response experiment on single species showed that the ER_{10} and ER_{50} of mecoprop-P were considerably lower for *C. bursa-pastoris* than for *G. dissectum*. Hence, it was a priori expected that *G. dissectum* would outcompete *C. bursa-pastoris* with increasing herbicide doses. However, this

expectation was not met in the competition experiment, even though the experiments were carried out at the same time and in the same greenhouse. Using a regression-based analysis, Damgaard et al. (2008) showed interactions between herbicide treatment and competition, and they propose that size-asymmetric competition may be an important factor, i.e. *C. bursa-pastoris* may have kept a dominating position in the competition for light even though this species was more affected by the herbicide in the single-species test. Based on these findings, they questioned the relevance of using single-species tests for the ecological risk assessment of herbicides.

A second experiment with two perennial grasses, *Festuca ovina* and *Agrostis capillaris*, and the application of glyphosate in a single-species standard test and in a two-species competition experiment with a complete additive design supported the finding by Damgaard et al. (2008) (Strandberg et al., 2007; Strandberg et al. submitted). For the competition experiment, seeds of the two grasses were sown in plant trays and at the two-leaf stage transplanted into polystyrene boxes (40 × 40 × 15 cm) filled with a potting mixture consisting of soil, sand and peat (2:1:1 by weight). A factorial design including 26 different combinations of densities of the two species was used. Each plant species was grown in monoculture at six densities (1, 4, 8, 16, 32 and 64 plants per box equivalent to 6 up to 400 plants per m^2), while binary mixtures were established in 14 boxes at densities from 8 to 64 plants per box. Glyphosate was applied at two doses, 22 and 61 g a.i./ha, equivalent to 1 and 5 % of the recommended dose, respectively, in a spraying chamber equipped with two ISO-02 nozzles operating at a pressure of 3 bars and a velocity of 5.6 km/h when plants were at the four- to six-tiller stage. Each treatment was replicated three times. Aboveground biomass was harvested six weeks after herbicide treatment. The plants in the outer 5 cm of the boxes were not included in the analyses. The fresh weight and number of plants were recorded. The plants were dried in an oven at 80 °C for 24 hours and dry weight was measured. It was shown that the herbicide increased the sensitivity of the most sensitive species, i.e. *A. capillaris*, by reducing the ER_{10} values estimated in the single-species test (19 g a.i./ha) by 16 %. Additionally, Strandberg et al. (2007, 2012) found that the outcome of the competition experiment corresponded with observed biomass data for these two species in a multi-species field experiment with the addition of low doses of glyphosate. A thorough description of the design of the field experiment can be found at: http://bios.au.dk/forskning/faciliteter/long-term-experimental-plot/

6.1.2.2. Microcosm experiments in the greenhouse for assessment of species interactions

Reuter and Siemoneit-Gast (2007) performed a comparative study of herbicide sensitivity of six plant species (*Leontodon hispidus*, *Silene nutans*, *Trifolium pratense*, *Galium mollugo*, *Bromus erectus*, *Cynosurus cristatus*) in single-species tests and in microcosms. The plants were exposed to two herbicides, one non-selective, i.e. Roundup Ultra (360 g glyphosate/L), and one herbicide with a narrower spectrum (selective against dicotyledons), i.e. Monitor (800 g sulfosulfuron/kg), at the two- to four-leaf stage. Plant trays (0.17m × 0.17m) were used for the microcosm experiment and each species was sown eight times in a uniform arrangement with a 2.5 cm distance. The 24 plants (four plants of each species) in the middle of the trays were sampled. The monocultures and microcosms were replicated four and three times, respectively. Aboveground biomass was assessed three times with 14-day intervals over the 42-day test period. In the single-species test, *B. erectus* and *T. pratense* were less sensitive than the other species, especially to the non-selective herbicide Monitor. Three species (*G. mollugo*, *L. hispidus*, *S. nutans*) were more sensitive to glyphosate in microcosms, whereas *G. mollugo* and *L. hispidus* were more sensitive to sulfosulfuron when grown in microcosms. *S. nutans* and *L. hispidus*, in particular, showed increased sensitivity in the microcosms having a sensitivity three times higher in microcosms than in monocultures. Recovery was also measured and only a few species showed recovery 42 days after treatment, and for *G. mollugo* (treated with both herbicides) and *S. nutans* (treated with sulfosulfuron) the sensitivity increased with time. Generally, phytotoxicity tests following OECD guidelines have a duration of 21–28 days and therefore may underestimate the sensitivity of some species. The authors generally concluded that effects measured with a few species in microcosms may not easily be transferred to the field situation.

Riemens et al. (2008, 2009) also performed comparative dose–response experiments of single-species tests and microcosms under standardised greenhouse conditions. For the microcosms, 5-L pots were used. Each microcosm consisted of four monocots (*Poa annua*, *Echinochloa crus-galli*, *Elymus repens*, *Panicum miliaceum*) and four dicots (*Solanum nigrum*, *Stellaria media*, *Chenopodium album*, *Centaurea cyanus*) that were placed alternately in the pots and thinned to eight plants per species per pot. Seeds were seeded in such a manner that emergence of the species would coincide. Four weeks later, the microcosms were sprayed with different rates (from approximately 1 to 100 % label rate depending on the herbicide and intended use) of the herbicide glufosinate ammonium in a spray chamber. Five herbicide rates and one control with eight replicates each were used. The first visual symptoms of herbicides were recorded two days after treatment, and, four weeks later, the fresh weight of the plants was determined. Although glufosinate ammonium is a non-selective herbicide, dicots were more sensitive than monocots and ER_{50} values of all monocots were higher than the ER_{50} values of the dicots. Additionally, Riemens et al. (2008) performed single-species tests with four species used in the microcosm experiments (*C. album*, *S. media*, *P. annua*, *E. crus-galli*) under greenhouse and field conditions. The results show that the greenhouse-grown plants had lower ER_{50} values than the field-grown plants, which might be a result of the different environmental conditions. Compared with the dicots, the monocots were also less affected in the microcosms. Although the study showed that the sensitivity of species grown individually and in mixtures differs from each other, it was not possible to separate inter- and intraspecific competition and shielding effects within this microcosm approach.

The study by Dalton and Boutin (2010) is an example of a comparative dose–response experiment that, in addition to comparisons of effects in single-species tests and microcosms, include comparisons of the importance of test conditions, i.e. greenhouse versus outdoor, for species sensitivity. The study used nine terrestrial (*Alliaria petiolata*, *Euthamia graminifolia*, *Fragaria virginiana*, *Geum canadense*, *Leucanthemum vulgare*, *Rudbeckia hirta*, *Solidago rugosa*, *Symphyotrichum lateriflorum*, *Symphyotrichum novae-angliae*) and seven wetland species (*Asclepias incarnata*, *Chelone glabra*, *Eupatorium maculatum*, *Eupatorium perfoliatum*, *Lycopus americanus*, *Phalaris arundinacea*, *Verbena hastata*) exposed to glyphosate and atrazine in separate experiments. The objective was to compare the response of the plants when grown individually in pots versus in different microcosms including 28 days after treatment tests in greenhouse and outdoor and long-term, i.e. 60–70 days after treatment, greenhouse tests. A microcosm consisted of a 5-L pot with one plant (*A. petiolata* for terrestrial and the grass species *P. arundinacea* for wetland microcosms) planted in the middle of the pot and the others planted in a circular arrangement around the perimeter of the pot. The plants were sprayed with a track sprayer in a spraying chamber when they had reached the four- to six-leaf stage. Generally, the greenhouse microcosms were the most sensitive tests with the largest biomass reductions. Plants in the single-species tests showed a similar sensitivity to plants grown for an extended test period or outdoors. They concluded that it is not possible to predict changes in community structure by single-species test.

Generally, the number of species used in the microcosm experiments was in a similar range (six to nine species). However, the individuals per species and per microcosm differed strongly between the three experiments. Dalton and Boutin (2010) used seven wetland or nine terrestrial species with only one individual of each species per microcosm. In contrast, Riemens et al. (2008, 2009) used eight different species and eight individual plants of each species per microcosm. Since both microcosm experiments used 5-L pots as test units, the plant density differed considerably between these two microcosm experiments. Generally, higher plant densities increase interactions between plants (i.e. competition) but shielding may also occur. Riemens et al. (2008), for example, detected shielding effects for the small species *Stellaria media*. This species probably received less of the applied herbicide when grown in microcosms owing to the shelter provided by other species. The microcosms in Reuter and Siemoneit-Gast (2007) differed slightly from the test system of Dalton and Boutin (2010) and Riemens et al. (2008), although the size of the test units corresponded to the 5-L pots.

The studies included both annual weeds and NTTPs as well as perennial NTTPs. Dalton and Boutin (2010) primarily used dicots (*P. arundinacea* was the only grass species included), whereas Riemens

et al. (2008, 2009) and Reuter and Siemoneit-Gast (2007) used a mix of dicots and monocots. Generally, a mixture of broad-leaved species and grasses seems to be a good choice, since most herbicides have a specific mode of action, targeting specifically mono- or dicots (Riemens et al., 2008) and even the non-selective herbicides do not affect all species equally. Moreover, in order to resemble the herbaceous vegetation of field margins and other semi-natural habitats, the microcosms should consist of such a mixture.

Herbicide applications in all studies presented here, including both greenhouse studies and the microcosms that were placed outdoors following treatment, were performed as overspray with drift-relevant herbicide rates, i.e. simulating drift. In all studies, the plants were treated a few weeks after emergence, at the four- to six-leaf stage (Reuter and Siemoneit-Gast, 2007; Riemens et al., 2008, 2009) or when they reached a size comparable to the four- to six-leaf stage (Dalton and Boutin, 2010) and biomass was used as an endpoint. Effects on reproduction were not investigated.

Test durations of 21–28 days as used in current standard plant testing may underestimate the risk of herbicides on plants. Reuter and Siemoneit-Gast (2007), for example, showed that the sensitivity of some species increased over time. In contrast, Dalton and Boutin (2010) found no increase in sensitivity in long-term (70–90 days after treatment) microcosm experiments.

Specifications and recommendations for greenhouse studies for assessment of species interactions

- No standard test guidelines exist for the assessment of species interactions. However, the above presented studies all include test designs that are suitable for assessment of the importance of species interactions.
- Schmitz et al. (2013b) provided a list of general recommendations on, for example, species selection, number of species, plant densities and test duration for microcosm studies.
- Generally, these tests are either very simplistic (two-species competition test) and hence easy to interpret or more realistic (more like natural plant communities) and therefore difficult to interpret. For example, it might not be possible to separate inter- and intraspecific competition and shielding effects. However, in the greenhouse, it is possible to control most environmental variables.

6.1.3. Comparing species sensitivity in greenhouse and under outdoor/field conditions

Conditions such as temperature, humidity and nutrient availability under which a test is conducted potentially affect the outcome of the test. Comparisons of species sensitivity to herbicides under standardised conditions in the greenhouse and under outdoor/field conditions, therefore, are an important focus in higher tier assessment.

Dalton and Boutin (2010) and Riemens et al. (2008) performed studies that, in addition to the comparison of species sensitivity in single-species test and microcosms (described above), encompass comparisons of species sensitivity in greenhouse and outdoor conditions, as described in the previous section.

Dalton and Boutin (2010) found that sensitivity was dependent on interactions between species and test conditions. For example, plants grown in the greenhouse were taller, greener and had more leaves than outdoor plants. Additionally, the temperature in the greenhouse was higher. These factors might have increased the translocation of the herbicide in the greenhouse plants and thereby increased their sensitivity. The outdoor plants had smaller leaves and presumably also thicker cuticles, which may have contributed to a decreased herbicide uptake, resulting in lower herbicide toxicity (Dalton and Boutin, 2010).

Riemens et al. (2008) found a linear relationship between the ER values of greenhouse- and field-grown plants treated with glufosinate ammonium, with greenhouse plants being more sensitive than

field-grown plants. The authors explained the discrepancy by the differences in environmental conditions, including temperature, relative humidity and light intensity. Riemens et al. (2008) described that a high relative humidity increases the efficacy of glufosinate ammonium due to hydration of the cuticle and, thus, a water-soluble compound such as glyfosinate ammonium can penetrate the cuticle more easily. A low relative humidity in the field results in a reduced uptake of the herbicide (Riemens et al., 2008).

Gove et al. (2007) performed a study that combines short-term greenhouse tests and long-term field experiments to investigate the effects of spray drift and fertiliser overspread on woodland ground flora. They exposed six-week-old herbaceous woodland species (*Mercurialis perennis*, *Primula vulgaris*, *Galium odoratum*, *Viola riviniana*, *Carex remota*, *Geranium robertianum*) potted separately and cultivated in a greenhouse to different rates of glyphosate (1, 5, 10 and 25 % of the field application rate (360 g a.i/L)). Then, half of the test plants remained in the greenhouse and the other half were transplanted into twenty 1-m^2 plots on two rows (10 plots on each row) at 6–7 m and 8–9 m from the woodland/field margin. Every plot contained one replicate of each herbicide treatment for all six species, i.e. 30 plants (five herbicide treatments × six species) per plot. The plants were randomly allocated in a grid position. Before, the plants were introduced, the plots were cleared and fenced and weeding was carried out over the course of the experiment to remove any other plants. Additionally, half of the plots were treated with a pelleted NPK (14-13-13) fertiliser with one rate equivalent to 50 % of the application rate for wheat (140 kg N/ha). One year later, the number of flowers per plant was recorded and the plants were harvested and weighted. The plants that remained in the greenhouse were also treated with the same fertiliser as used in the field or with distilled water (control). Each treatment was replicated 10 times and the plants were harvested 10 weeks later. Gove et al. (2007) found good agreement between the short-term greenhouse and long-term field experiments, although long-term impacts may be underestimated in the short-term experiment. The results showed an increased mortality, reduced biomass and reduced fecundity for all six species treated with glyphosate compared with the control. Glyphosate application rates as low as 5 % resulted in a reduction in the proportion of flowering plants. In contrast, the fertiliser treatment affected resource portioning in *C. remota* and *G. odoratum* and reduced the fecundity of *G. odoratum*.

Pfleeger et al. (2012) aimed at developing a simple tier-III field test that was economical, was geographically flexible, used relevant test species and ecologically relevant endpoints and compared results with the standard single-species test in the greenhouse. Three native plant species (*Festuca roemeri*, *Clarkia amoena*, *Prunella vulgaris*) and one introduced species (*Cynosurus echinatus*) were exposed to glyphosate and aminopyralid. The experiment was replicated at two locations and repeated for two and three years for aminopyralid and glyphosate, respectively. The individual test plots were 60 cm × 60 cm. Plant height, width and volume were used as endpoints and measured every second week. With glyphosate, the relative rank in species sensitivity among the species differed between the greenhouse and field, with *C. echinatus* being the most sensitive species in the field and *P. vulgaris* the most sensitive in the greenhouse. With aminopyralid, the ranking of all species based on the sensitivity was similar for the greenhouse and field. Based on these results, the authors concluded that ecological effects of herbicides on plant communities can be investigated by the proposed simple test design, although interactions between species among others are not included and only vegetative endpoints are included.

6.1.3.1. Specifications and recommendations for studies comparing species sensitivity in the greenhouse and under outdoor/field conditions

No standard test guideline exists for the assessment of the importance of test conditions (greenhouse vs. outdoor conditions) for species sensitivity. However, the above presented studies all include test designs that are suitable for the assessment of the importance of test conditions. Generally, the studies compare the effects assessed in single species/microcosms in the greenhouse with effects assessed under outdoor conditions.

6.1.4. Phytometer experiments

Several studies mention that exposure routes, e.g. overspray/direct spray and spray drift, and the differences between these with respect to, for example, droplet size and herbicide concentrations may affect the measured effects on NTTPs (Koch et al., 2004; De Snoo et al., 2005; Strandberg et al., 2012). Drift consists of smaller droplets with possible higher concentrations of the pesticide than overspray. Furthermore, drift in the field is very sensitive to meteorological conditions (e.g. wind speed, temperature, relative humidity) and technical factors (e.g. boom height, driving speed, nozzles) and these factors can vary from application to application and may produce different effects. The main advantage of direct spray is that the application can be performed under controlled, repeatable conditions, and the spray volume can be kept constant. However, no study has been able to conclude whether these differences in exposure produce different effects on the plants (Strandberg et al., 2012).

One way to study the effects of actual spray drift on individual species or microcosms is by phytometer experiments, where the plants/microcosms are placed in the field and are exposed to drifting herbicides applied on the neighbouring field. The main purpose of these experiments, however, has been to detect the in-field buffer distances to protect the vegetation of field margins. Schmitz et al. (2013b) provide a thorough summary.

Marrs and co-workers conducted a series of tests to investigate the effects of herbicide drift on native plant species of conservation interest (Marrs et al., 1989, 1991a, 1993; Marrs and Frost, 1997). They used single plants (annual weeds and NTTPs), as well as microcosms composed of eight dicots and eventually a grass, for the experiments. Each microcosm (pot: 27 cm diameter × 12 cm deep) contained one individual plant of each species. The plants/microcosms were placed at different distances downwind of the field of application. Some of the studies, i.e. Marrs et al. (1989) and Marrs and Frost (1997), included observations over several years (two to four years), which made it possible to include effects of repeated herbicide exposures in the assessments. An additional advantage of study periods lasting more than one year/season is that effects on reproduction in all species, including biennials and perennials, can be quantified (Marrs et al., 1989).

The first study included 23 non-crop plant species and five herbicides: asulam, chlorosulfuron and metsulfuron methyl, glyphosate, MCPA and mecoprop (Marrs et al., 1989). The assessments showed that lethal effects were present up to 6 m away from the treated field. Following Rautmann et al. (2001), the estimated drift at this distance from the field is 0.48 % of the field rate. Effects on flowering, i.e. absence of flowering, and seed production were found up to 10 m from the field and the greatest distance at which damage effects (e.g. reduction in size, leaf chlorosis, discoloration) were found was 20 m (0.15 % of the field rate). In general, Marrs et al. (1989) found that some species appeared to be consistently more sensitive than others, e.g. *Cardamine pratensis*, *Centaurea nigra*, *Digitalis purpurea*, *Lychnis flos-cuculi*, *Medicago lupulina* and *Prunella vulgaris*. Based on this study, they suggested in-field buffer zones of 5–10 m for ground applications to minimise the risk of herbicide impacts on the vegetation of field margins.

In two subsequent studies, Marrs et al. (1991, 1993) focused on the importance of plant age/physiological stage at the time of exposure. These studies confirmed that young plants were more affected than older ones when using survival and biomass reductions as endpoints, and concluded that buffer zones for established plants could be set at 6–10 m, but, where seedling regeneration is important, a buffer zone of 20 m is needed.

Although there are a number of microcosm studies, Marrs and Frost (1997) were only able to make a few generalisations on the effects of herbicides on plants when they are growing in mixtures and concluded that "we are a very long way from being able to predict the outcome of spray events on multi-species mixtures". The main result is that the response of plants in mixtures to spray drift depends on the herbicide and the species present, and in particular whether there are grasses among them.

The study of de Jong and de Haes (2001) aimed at developing a sound test procedure for assessing the short-term impact of low rates of herbicides on vascular plants using an iterative procedure. At first they tested various designs in the greenhouse, and subsequently used the best method for the field experiments. Experiments were conducted with three herbicides (glyphosate, bentazone and diquat) using *Brassica napus* and *Poa annua* as test species. For *B. napus*, 150 seeds were placed in a multi-compartment tray (30 × 50 cm, 10 × 15 compartments) and, for *P. annua*, a 10-L plant box was used, which was divided into three parts. In each part of the box, 0.075 g *P. annua* seeds were sown. The plants were grown in the laboratory for approximately two weeks and then transferred to the field for spraying. The trays and boxes were placed at distances of 0, 2, 4, 8 and 16 m downwind of the sprayer, and control trays were placed > 500 m from the treated area. The test plot was sprayed with a knapsack sprayer, which was connected to a 1-m spray boom. Two hours after spraying, the trays and boxes were returned to the cultivation rooms. Per treatment, 50 *B. napus* plants were harvested 7, 14 and 21 days after treatment; biomass per individual was measured for 20 randomly chosen plants, and survival and total biomass was recorded for the remaining 30 plants. For *P. annua*, one measurement was performed on 30 plants to get an accurately measurable amount, since individual plants of *P. annua* were very small. Additionally, the deposition rate of the applied spray volume was determined using water-sensitive paper. Considerable differences in biomass and growth were detected between species as well as between herbicides. For glyphosate, the distance at which 50 % growth inhibition occurred for both species was between 5 and 6 m from the test plot, equal to 0.57 and 0.48 % of the field rate. Diquat led to a 50 % growth inhibition at this distance only in *B. napus*. However, in some of the experiments, a 50 % growth inhibition was found 16 m from the sprayed area. de Jong and de Haes (2001) concluded that these bioassay tests were suitable for the assessment of impacts of herbicide drift on plants.

6.1.4.1. Specifications and recommendations for phytometer experiments

No standard test guideline exists for the assessment of the importance of exposure routes for the effects of herbicides on NTTPs. The phytometer experiments described above, however, represent some useful methods.

6.1.5. Field experiments with experimental exposure of existing vegetation

There have been very few studies of the effects of herbicides on NTTPs established in the field that include an experimental approach applicable for higher tier testing. However, these studies form an important contribution to the study of higher tier effects, as they suggest methods for evaluating herbicide effects on shrubs.

Kjær et al. (2006a) investigated the effects of overspray of hawthorn (*Crataegus monogyna*) with metsulfuron at doses simulating spray drift. Hawthorn is a common shrub in hedgerows in agricultural areas and around orchards. Kjær et al. (2006b) performed a fully randomised spray experiment in seven hawthorn hedgerows with four doses of the herbicide, equal to 5–40 % of the field rate, and a control. Spraying was done at the bud stage and at early flowering, and the number and biomass of leaves, flowers, green berries and mature berries were recorded. The results showed that hawthorn was most sensitive when it received a spray application at the bud stage. Spraying at this stage caused a highly significant reduction in the number and dry weight of berries, whereas it had no effects on leaf and flower production. A 100 % berry reduction was found with herbicide doses of 5 % of the label rate. Spraying at early flower stage also reduced the number of berries significantly, although to a lesser extent. The dramatic reductions in berry production may have serious consequences for berry-eating birds, including numbers of nesting birds. The year after the application, shrubs were revisited and the effects of the herbicide drift on the same endpoints as the year before were measured (Kjær et al., 2006b). This revealed significant effects on both growth indices and reproductive endpoints. Based on these results, the authors concluded that, at present, the effects on non-target plants are likely to be underestimated since the risk assessment focuses on results from short-term laboratory studies and the effects on reproduction are not assessed.

6.1.6. Field experiments using experimental established vegetation

In recent years, an increasing number of studies of herbicide effects on plants have used experimental approaches that form an important contribution for higher tier assessment of herbicide effects on NTTPs.

In 1996, Perry et al. established an experiment to determine the effects of herbicide spray drift (glyphosate) and fertiliser (ammonium nitrate fertiliser with 34.5 % N) on a simulated field margin community containing three grasses (*Elymus repens*, *Arrhenatherum elatius*, *Bromus sterilis*) and three broad-leaved plants (*Ranunculus repens*, *Silene latifolia*, *Galium aparine*) (Perry et al., 1996). The experiment was laid out as a randomised block design with four replicated blocks each containing 12 treatments, i.e. three fertiliser treatments (0, 50, 200 kg N/ha) and four herbicide rates (0, 45, 90, 180 g a.i./ha). Each plot was 2 × 3 m and plots were separated by 70 cm. Plots were hand weeded the first year of establishment to prevent invasion by other species. The first fertiliser treatment was carried out in March 1995, i.e. 11 months after the establishment the plots, and in June 1995 the plots were treated with the herbicide. The monitoring of the plots started in March 1995. Plant cover was measured by the point-intercept method using a 1 m high point quadrat frame, which contained 10 pins. In each plot, the frame was randomly positioned three times and the numbers of touches of each species on each pin were recorded at height intervals of 5 cm. The assessments were repeated monthly from March to August. At the beginning of the assessment there were no visible differences between the plots. However, *R. repens* and *G. aparine* failed to establish, presumably because no pre-seeding treatment was carried out and therefore these two species were not taken into further consideration. The results showed that fertiliser and herbicide applications had a significant effect on the four established plant species. All fertiliser treatments caused a significant reduction in cover of *S. latifolia* and *A. elatius* and all rates of glyphosate significantly reduced the cover of the sown grasses. These effects became stronger with time. Interaction effects between the fertiliser and herbicide treatments were not found. The authors supposed that this could change with time (Perry et al., 1996). However, the experiment has not been reported since.

Kleijn and Snoeijing (1997) assessed the botanical changes caused by low levels of herbicide (fluroxypyr) and fertiliser. The experiment was established in 1993 on low productive grassland dominated by *Festuca rubra* ssp. *commutata* and *Holcus lanatus*, and the normal management practices including autumn cutting and removal of vegetation were maintained. The experiment was set up as a randomised complete block design with four replicates of the 12 treatments (three fertiliser levels (0, 27.5, 55 kg N/ha) × four herbicide levels (0, 10, 20, 100 g a.i./ha)). In total, the experiment included 48 plots, each 2 × 2 m and 0.5 m apart. The fertiliser and herbicide were applied simultaneously once a year in spring for three consecutive years. Vegetation composition was recorded once a year in early spring, i.e. before treatment, in the central square metre, leaving a buffer of 0.5 m on all sides. Aboveground biomass was harvested in two 0.3 × 0.3 m subplots within every plot in late August every year. Generally, fertiliser was found to have larger effects on both species richness and biomass than herbicide. Significant herbicide effects were mainly limited to the highest dose.

In 2001, a long-term experimental plot was established to investigate the combined effects of low doses of herbicide (glyphosate) and nitrogen on grassland communities (http://bios.au.dk/forskning/faciliteter/long-term-experimental-plot/). Before sowing, the area was ploughed to 60 cm. In spring 2001, 31 species were sown. The species selected were grassland species covering different life form strategies (Grime, 2001). The experimental manipulations were set up as a complete randomised block design with 10 replicates of each of the 12 treatments, including four glyphosate treatments (0, 14.4, 72 and 360 g a.i./ha equal to 0, 1, 5 and 25 % of label rate of 1 440 g glyphosate/ha, respectively) and three nitrogen treatments (0, 25 and 100 kg N/ha). Additionally, all plots received phosphorus (53 kg/ha), potassium (14 kg/ha), sulphur (50 kg/ha) and copper (0.7 kg/ha) every year. The RoundupBio® formulation of glyphosate was used for the experiment. For the herbicide applications, a 3-m beam with 0.5 m between the nozzles (Lurmark Lo-drift LD 015 Green nozzles with a pressure of 2.0 bars) was used, and fertilisers were spread by hand. The plots were

treated with glyphosate for the first time on 24 August 2001. Since then, it has been treated with herbicide and fertiliser once every year in spring (mid-May). Every March, woody species were removed both to keep the area as grassland and to allow the use of the spraying equipment. Since 2005, plant cover has been sampled at least once a year, but in most years three samplings are carried out: pre-herbicide treatment (mid-May), two weeks after treatment (mid-June) and in mid-August. Plant cover was estimated within six randomly selected 0.75 m × 0.75 m quadrats by the pin-point (or point-intercept) method using a horizontal frame with a 5 × 5 grid with the 25 intersections at a distance of 10 cm. At each intersection, a sharply pointed pin with a diameter of 0.5 mm was passed vertically through the vegetation. An estimate of percentage cover of vascular plants was obtained by recording the first interception of the pin with the canopy of the different species or ground.

Generally, Strandberg et al. (2012) found that applications of glyphosate at spray drift-relevant doses have resulted in decreased species diversity compared with controls, and the results also showed interactions between species with different sensitivities to glyphosate. Both nitrogen and glyphosate treatments had significant effects on the species composition and cover of the nine most abundant plant species, including four grasses, i.e. *Festuca ovina*, *Elytrigia repens*, *Agrostis gigantea*, *Agrostis capillaris*, and five dicotyledons, *Tanacetum vulgare*, *Euphorbia esula*, *Leucanthemum vulgare*, *Hierachium pilosella* and *Linaria vulgaris*, and this varied significantly with treatment. Generally, species cover decreased with increasing glyphosate doses. The cover of *F. ovina* and *E. esula* were the only exceptions. For these species, the cover increased with increasing dose of glyphosate. Although the cover of *F. ovina* was high in 2005 and has continued to increase in plots treated with glyphosate, the cover of *E. esual* was low in 2005 and has increased over the years. Increasing nitrogen generally resulted in increasing plant cover, except in species adapted to nutrient-poor conditions, i.e. *F. ovina* and *H. pilosella*.

Schmitz et al. (2013b) offered suggestions for the design and performance of tier III field studies using four field tests designed to evaluate the effects on plant communities. They suggested that experimental study sites should not be contaminated with agrochemicals (pesticides, fertiliser). An appropriate study site would be a meadow with a relative homogenous distribution of approximately 40–50 different plant species. Such a meadow can be regarded as an original habitat that is not contaminated with agrochemicals and therefore represents the plant community of a surrogate field margin. The design of a field experiment and its statistical analysis are intimately connected. Therefore, the experimental test design has to be well described and has to take into account potential underlying environmental gradients. An appropriate test design would be, for example, a randomised block design. The size of the test plots is dependent on the number of species of the study sites and the homogenous distribution of these species. However, the size of the test plots should be not too small. The number of replicates in the evaluated field studies should range between 4 and 14. Since community analyses are complex, it is important to increase replication whenever possible (Fraser and Keddy, 1997). The time and number of applications should be in agreement with label recommendations. The herbicide product should be applied, not just the active ingredient. The test duration should be considerably longer than a tier II study, since effects on reproduction and plant composition should be investigated. These effects are often only apparent in the next growing season (one year after treatment).

There are no regulations for fertiliser applications next to field margins. Therefore, it seems necessary to consider the nutrient inputs on plant communities and their interactions with herbicides as well. The vegetation of the study site should be assessed before and after treatments at different time intervals. Plant community assessments have to be performed with a method that is appropriate in order to document changes in the plots over time. In addition, it is important to use a method with which uniform plant community assessments can be obtained, independent of the technicians. At the end of the growing season, biomass samples from each plot (e.g. aboveground biomass of 1 m × 1m) should be taken and measured. Effects on reproduction (flowering, seed set) should also be recorded.

On grass-dominated meadows, Schmitz et al. (2013a) established a field experiment to study the single and combined effects of repeated herbicide, insecticide and fertiliser applications on *Ranunculus acris*

in successive growing seasons. In total, the meadow flora included approximately 40 herbaceous plants and 13 grasses. The experimental treatment was set up as a randomised block design with eight replicates of each of the three treatments (herbicide, insecticide and fertiliser). Each plot was 8 m × 8 m with a 2-m distance between plots. Atlantis WG (sulfonylurea) was used and applied once a year in April. The local management practice for field margins, with cutting and subsequent removal of the cut vegetation in early July, was maintained during the experiment. *R. acris* started to sprout early in the spring and at the time of herbicide application its phenological stage was one to two weeks before the onset of flowering. The application of both herbicide and fertiliser decreased *R. acris* plant density significantly and, in addition, the herbicide caused an 85 % reduction in its flower intensity.

Guidelines for field testing were made available to registrants by the US EPA in 2012: Ecological Effects Test Guidelines: Terrestrial Plants Field Study (US EPA, 2012d). The intention of these guidelines is to describe general procedures for performing plant toxicity tests under field conditions, both in field and off field. The purpose of the field study is to quantify the risk that may occur to terrestrial plants, plant populations or plant communities from pesticide use. The guideline provides information on factors to be considered in the design and conduct of field studies for effects of chemical substances and mixtures on terrestrial plants. The timing of applications, test conditions, selected site characteristics and habitats, number of sites, types and number of plant species tested, geographical areas and experimental design are dependent on the questions and are decided on a case-by-case basis. Effects considered may include effects at the individual plant level (mortality, sublethal toxic effects (such as decreased biomass) or other morphological changes), or changes in population or community parameters. In this guideline, a community is defined as an assemblage of populations of different species and a population is defined as a group of individuals of the same species. Effects on plant reproduction measured through the production of flowers, pods, fruits or seeds or viability of seeds may be included. This US EPA guideline should be used in conjunction with a background document which provides general information and overall guidance on test procedures, equipment, statistical analyses and reporting.

6.1.6.1. Specifications and recommendations for field experiments

No standard test guideline exists for field assessment of the effects of herbicides. To date, no standard protocol is available for conducting field tests appropriate for tier III studies. However, two sources of recommendations for field studies are available: general recommendations on field studies by US EPA (2012d) and recommendations for tier III field studies by Schmitz et al. (2013b).

6.1.7. Herbicidal effects on ecosystem function and services

Herbicides have the potential to affect ecosystem functions and properties through their influence on the species present and their abundance. Both theory (the "mass ratio hypothesis") and experimental evidence suggest that the extent to which a plant affects ecosystem function is likely to be predictable from its contribution to the total biomass. As such, Grime (1998) concluded that functional diversity among dominants and perhaps also within subordinates, i.e. a species forming a lower proportion of the biomass, but sometimes more numerous as individuals than dominants, is capable of having an immediate impact on the properties of ecosystems.

A few studies have looked at vegetation changes along gradients of herbicide exposure (Gove et al., 2007; Boutin et al., 2014) or have compared vegetation within habitats adjacent to organic fields, with similar habitats adjacent to conventional fields receiving pesticides on a regular basis (Boutin et al., 2014; Strandberg et al., unpublished data). All studies found increased species diversity with decreased/no herbicide exposure and Damgaard et al. (accepted) showed increasing functional diversity of hedgerow ground vegetation with increasing number of years since the transition to organic management of the neighbouring field. Strandberg et al. (unpublished data), however, showed that the same species made up the majority of the plant cover of hedgerow ground vegetation adjacent to organic or conventional fields, but an increasing number of less abundant species were found with increasing number of years since the transition to organic farming. This indicates that organic farming, and presumably also other instruments that eliminate or decrease herbicide exposure, may lead to an

increasing complexity of the vegetation that may, for example, sustain a more diverse pollinator and predator insect fauna with the potential to provide the neighbouring field with important ecological services (Petersen et al., 2006; Jonason et al., 2011). This positive effect of conversion to organic farming, i.e. no crop pesticide treatments, on pollinators and other flower-visiting insects is further enhanced, as plants in organic hedgerows have significantly more flowers, start flowering earlier and flower for a longer period than the same species do in hedgerows adjacent to conventionally herbicide-treated fields (Boutin et al., 2014).

6.2. Population modelling

6.2.1. Models for estimation of plant interactions based on observations of changes in density, biomass or plant cover

Several models are available for studies of the impact of plant interactions/plant competition on plant population and community dynamics based on observations of changes in density, biomass or plant cover. These include two models restricted to the analysis of interactions of annual plants based on time series data: (1) statistical methods for the analysis of spatially structured population data (number of flowering individuals, fruit numbers and number of seeds in the seed bank) on four species of winter annuals in coastal habitats for a 10-year period (Rees et al., 1996) and (2) a discrete hyperbolic competition model for analysis of species interactions of two competing genotypes of *Arabidopsis thaliana* that vary in sensitivity to the fungal pathogen *Peronospora parasitica* in an experimentally manipulated study with and without the fungal spores (Damgaard, 2003). Rees et al. (1996) demonstrated that interspecific interactions are extremely weak relative to intraspecific ones in the studied communities and that the spatial arrangement of species and individuals within them is critical to the observed dynamics. Damgaard (2003) discusses the potential for adapting the competition model to analyse data obtained in natural plant communities with an environmental gradient. Important to both models is that time series data are available and that the sampling area covers some variation in plant density and frequency among species.

As stressed above, often it is not possible to distinguish individual plants and determine plant density in many natural and semi-natural plant communities dominated by perennial plant species that form dense vegetation. For the analysis of species interactions in such communities, Damgaard (2011) developed a novel method that allows data on estimated plant cover and biomass of individual species sampled in permanent pin-point quadrats to be obtained.

6.2.2. Selection of species and endpoints for effect studies and modelling

Regardless of the study method, the selection of species, endpoints and spatio/temporal scales need to be considered. Some characteristics of plants are important to use or take into account in modelling. These include lifespan, plant size (height, root–shoot ratio), leaf size and shape, pollination strategy, seed production, seed dispersal, seed bank and size of populations. Many of the studies described in the previous sections have used annual plants, as this considerably simplifies the modelling. For plants with more complex life histories, a range of modelling approaches is available. Perhaps the most well-known approach is matrix modelling (Caswell, 2001). Here individuals are classified into a small number of states, for example small, medium and large individuals, and a transition matrix is used to project the population forward in time. Reproduction and the movements of surviving individuals between stages is governed by a population projection or Lefkovitch matrix, M. The dynamics are then given by:

$$\mathbf{n}(t+1) = \mathbf{M}\mathbf{n}(t)\,. \qquad (1)$$

The Perron–Frobenius theorem applies provided that M is power positive, so the long-term growth rate is given by the dominant eigenvalue, λ_1, of M, and the stable stage distribution by the corresponding eigenvector, w_1.

To give a concrete example, here is the projection matrix used by Shea and Kelly (1998) to explore the dynamics of *Carduus nutans,* an invasive thistle:

$$\mathbf{M} = \begin{array}{c} \\ SB \\ S \\ M \\ L \end{array} \begin{array}{c} \begin{array}{cccc} SB & S & M & L \end{array} \\ \begin{bmatrix} 0.04 & 8.25 & 179.41 & 503.14 \\ 0.19 & 1.09 & 22.18 & 62.18 \\ 0 & 0.01 & 0 & 0 \\ 0 & 0.01 & 0.02 & 0 \end{bmatrix} \end{array} \quad \text{.....(2)}$$

SB is the number of seeds in the seed bank and *S*, *M* and *L* refers to thistle rosettes that are small, medium and large in size. The matrix has the following simple interpretation: each column gives the expected contribution of a particular stage to each of the other stages. Therefore, the first column shows that 4 % of the seeds in the seed bank will stay there and 19 % will become small rosettes; the second column shows that each small rosette will give rise to 8.25 seeds in the seed bank, 1.09 small rosettes, and a small number of medium and large rosettes, and so on.

Constructing the matrix M for a real population requires selecting appropriate stages. If the life cycle is divided into discrete stages, this is straightforward. Otherwise things become more complicated, as it is necessary to (1) decide on the appropriate measure of individual state and (2) set the boundaries between stages. Practical issues of data collection and the ability to predict an individual's fate may determine how to measure an individual's state. Typically a single variable is used (e.g. longest leaf length or rosette diameter as a measure of plant size), but more complex classifications, say by age and size, are also possible. Setting boundaries may be problematic. Ideally there should be many categories, so that all individuals within a category really behave in a similar way, as the model assumes. However, the more categories there are, the fewer observations there are on each category, so estimates of the elements of M become less reliable.

Integral projection models provide an elegant way around these problems. To avoid this problem, in 2000, Michael R. Easterling, Stephen P. Ellner and Philip M. Dixon proposed the integral projection model (IPM), where individuals are characterised by a continuous variable *z* such as size (Easterling et al., 2000). The state of the population is given by $n(z,t)$, such that the number of individuals with sizes between *a* and *b* is

$$\int_a^b n(z,t)\,dz$$

Instead of the matrix M, the IPM has a projection kernel $K(z',z)$, so that

$$n(z',t+1) = \int_s^S K(z',z)n(z,t)dx$$

where *s* and *S* are the minimum and maximum possible sizes. The integration is the continuous version of equation 1, adding up all the contributions to size z' at time $t + 1$ by individuals of size *z* at time *t*. Providing some technical conditions are met (see Ellner and Rees (2006) for details), the IPM behaves essentially like a matrix model.

Both matrix and integral projection models require data on individual-level demography, that is marked individuals that are followed through time (Figure 3).

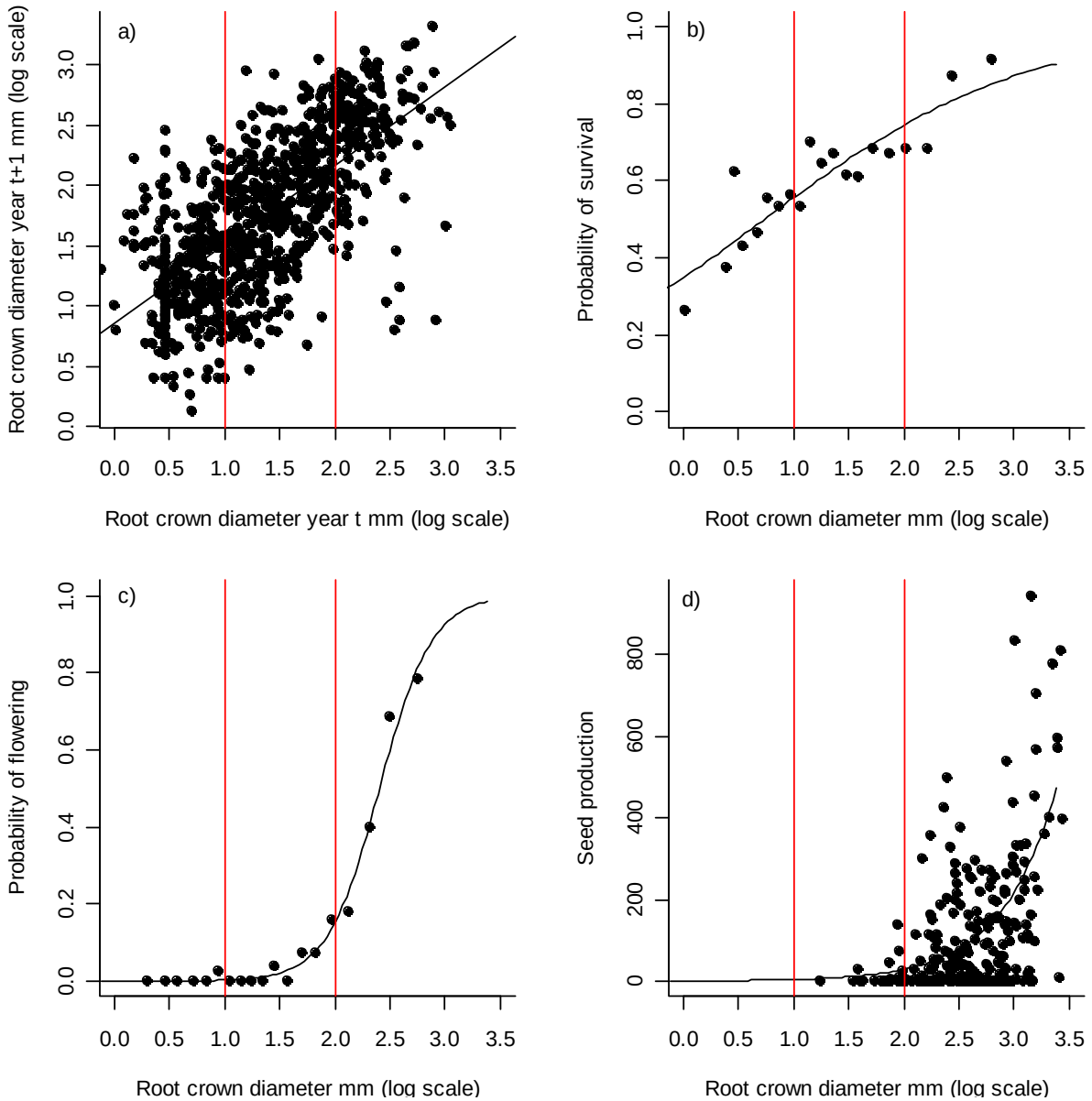


Figure 3: Size-structured demographic rates for Platte thistle, *Cirsium canescens*. (a) Growth (as characterised by plant size in successive years), (b) survival, (c) the probability of flowering and (d) seed production all vary continuously with size and can be described by simple regression models. Redrawn from Rose et al. (2005). In panels (b) and (c), the data were divided into 20 equal-sized categories and the plotted points are fractions within each category, but the logistic regression models (plotted as curves) were fitted to the binary values (e.g. flowering vs. not flowering) for each individual.

These models are constructed using regressions and so, in principle, it would be possible to parameterise a model including the effects of herbicide application. This would have, say, a different intercept if the growth model for plants that received herbicide, and then the consequences of this on population growth, could be determined.

One particularly important aspect of any modelling exercise is the relative timing of herbicide application and the action of density dependence. To illustrate this, consider the following simple model. We assume there are microsites in which seedlings compete such that, regardless of the

number of seedlings within a microsite, F seeds are produced per microsite. If seeds are Poisson distributed across microsites, then we have the following simple population model:

$$S_{t+1} = F(1 - \exp(-S_t))$$

where S_t is the density of seeds per microsite in year t. If we assume the fraction of seedlings that survive herbicide is ρ_b for herbicide application before density dependence acts and ρ_a for application after density dependence, the model then becomes

$$S_{t+1} = \rho_a F(1 - \exp(-\rho_b S_t))$$

where, for simplicity, we have assumed that the application of herbicide is always the same every year. We will explore three scenarios:

1. no herbicide application;
2. herbicide application before density dependence ($\rho_b = 0.2$, $\rho_a = 1$)($\rho_b = 0.2$, $\rho_a = 1$);
3. herbicide application after density dependence ($\rho_b = 1$, $\rho_a = 0.2$)($\rho_b = 1$, $\rho_a = 0.2$).

Therefore, 80 % of seedlings die after herbicide application. Solving numerically for the equilibrium the density of seeds per microsite allows the effects of herbicide drift to be explored (see Figure 4). When there is no herbicide drift, the equilibrium population size is ~F (black line in Figure 4). When herbicide application occurs before the action of density dependence, the equilibrium population size is reduced as expected and the largest reductions occur when plant fecundity is low (green line Figure 4). At high fecundity, the effects of herbicide drift are negligible, as there are large numbers of seedlings and many of those that are killed by herbicide application would have died as a result of density dependence. Hence, despite 80 % of all seedlings dying as a result of herbicide drift, there is very little impact on population size. In contrast, when herbicide drift occurs after density dependence, the impact on equilibrium population size is much greater (red line Figure 4). Clearly, the timing of herbicide application is critical for understanding the impact of herbicide drift on the equilibrium population size.

In both herbicide application cases the population goes extinct then $F\rho_a\rho_b < 1$,$F\rho_a\rho_b < 1$, which, for the cases analysed in Figure 4, means that all populations exposed to herbicide drift go extinct when $F < 5$. In this case, the timing of drift does not affect the condition for population persistence. This is a consequence of population persistence being determined by plant performance when rare, when there is no density dependence.

This model illustrates that the ecological assessment of herbicide drift depends on the timing of herbicide application relative to density dependence when considering the impact on population size but not on population persistence. The impact on population size is also determined by ecological conditions, for example in high-fertility situations (large fecundity) where herbicide drift occurs before density dependence and then populations are strongly buffered and, even with 80 % seedling mortality, the reduction in population size is negligible; however, there are large effects in low-fertility situations (low fecundity). These effects are in turn small relative to the reduction caused if herbicide drift occurs after density dependence. These effects mean that assessing the ecological impacts of drift using data from standard toxicity trials will be difficult, as the impacts are highly context specific. Adding layers of ecological complexity, for example by incorporating size-dependent demography as in the integral projection models discussed above, will only exacerbate this problem, as there will be multiple, potentially sublethal, effects on different components of demography (e.g. reductions in growth, survival and fecundity) and potentially each of these will be mediated by complex density-dependent processes. For example, herbicide-induced plant mortality might result in more rapid growth in those that survive, which in turn increases seed production, and offsets the impact of reduced plant survival.

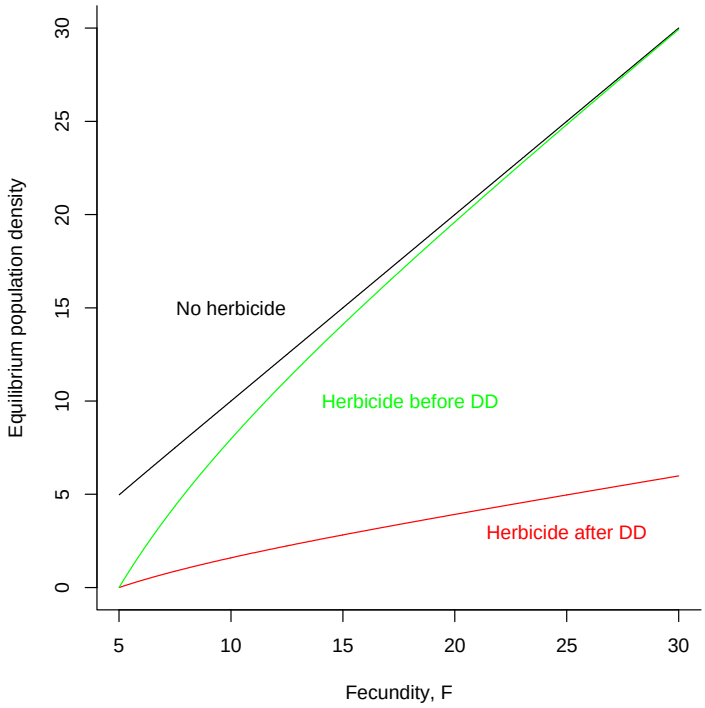


Figure 4: Effect of herbicide drift on equilibrium population size for the three scenarios discussed in the text.

6.2.2.1. Specifications and recommendations for selection of species and parameters important to measure

Extrapolate from individual effects to effects on populations and fitness

Consider possible extrapolation from visual effects to biomass, biomass to population or directly visual effects to population

- What are relevant species for which to develop models
- Think of biological input, model description, and appropriate scenarios
- Size of population/spatial scale (needed to sustain the population)
- Aim is to define SPGs and "validate" scheme
- What kinds of models are available?
- What are the questions addressed by the available models?
- Parameters model are built on, critical input parameters?
- Need for models that can deal with reproductive endpoints as input
- Models that can use endpoint of biomass as input and relate this to effects on the population?
- Dependence on the density of the plant under investigation, e.g. rare plants?

Characteristics of plants that are important to use or to take into account in modelling:

- lifespan;
- size of plants—height;
- leaf shape/area;
- pollination strategy;
- seed production;
- seed dispersal;
- seed bank;
- size of populations.

7. Comparison of outcome of lower and higher tier calibration to keep the same level of protection

7.1. Phytotoxicity studies using single-species and multi-species tests

Single-species tests with plants growing individually in pots or as monoculture are used for phytotoxicity assessment. They are conducted under ideal greenhouse conditions where wind, drought, competition, predators and other stressors are usually absent. Results of these tests are used to assess effects of herbicides on plant communities growing outdoors under natural conditions where a multitude of variable conditions may be present. Thus, two sets of confounding factors are involved: the level of plant competition and abiotic conditions. Not surprisingly, contradictory results have emerged from the few studies that have been performed comparing single-species tests in greenhouses and multi-species field experiments.

For agronomic purposes, many studies have been conducted to test herbicide efficacy first under greenhouse conditions in single-species tests followed by small plot testing in outdoor conditions. All these studies were conducted with high doses for the purpose of testing herbicide efficacy. In many instances, there was good agreement between greenhouse and field assessments. Sweat et al. (1998) examined the efficacy of 21 herbicides for the control of four *Amaranthus* spp. in soybean at the recommended label rate. *Amaranthus* species were grown in monoculture pots in greenhouses while they were sown with soybean in small plots outdoors. Although different experimental conditions prevailed, it was concluded that greenhouse results were similar to field experiments. Similar conclusions were reached in other studies comparing greenhouse single-species tests and small plot experiments (Jacques and Harvey, 1979; Glover and Shapaugh, 1997; de Jong and de Haes, 2001; McCalla et al., 2004; Richardson et al., 2004; Kegode and Fronning, 2005; Lingenfelter and Curran, 2007). On the contrary, several studies concluded that the response of plant species was more pronounced in single-species testing in growth chambers than in multi-species testing in fields (Friesen et al., 1976; Cornish and Burgin, 2005). Tan et al. (2002) found that effects were similar in greenhouses and field experiments except that it took longer under field conditions for effects to be manifested. Contrasting results were obtained in other studies between herbicides (Taylor-Lovell et al., 2001) and plant species tested (Mayo et al., 1995). There was either concordance between greenhouse and field studies (two herbicides) or more injury in greenhouse plants than in field plants in the case of one herbicide (Taylor-Lovell et al., 2001). Mayo et al. (1995) found that, in general, similar results were obtained in single-species greenhouses and small plot field experiments except that one species was more difficult to control under greenhouse conditions. Using the PHYTOTOX database encompassing several studies under variable conditions, Fletcher et al. (1990) revealed that 6 out of 20 species were more sensitive when tested in greenhouses than in field, whereas 11 out of 20 were more sensitive under field conditions.

Only a few studies have been performed with the explicit intention of comparing single-species and multi-species tests in dose–response experiments. Pfleeger et al. (2011) found that potatoes, peas and soybean plants tested with four herbicides in pots in the greenhouse and outdoors at two or three developmental stages generated comparable results. Conversely, four non-crop species (*Festuca*

roemeri, *Clarkia amonea*, *Prunella vulgaris* and *Cynosurus echinatus*) tested with two herbicides, glyphosate and aminopyralid, generated different results when tested in greenhouses and in the field (Pfleeger et al., 2012). Damgaard et al. (2008) tested *Capsella bursa-pastoris* and *Geranium dissectum* in the greenhouse, singly in pots or grown together with the herbicide mecoprop-P. It was concluded that single-species tests could not predict the outcome of effects when the two species were grown together in a competitive experiment. Kleijn and Snoeijing (1997) studied the effect of four doses of the herbicide fluroxypyr on the seedling stage of 18 species grown individually in pots in a greenhouse. The same species were sown in a mixture under field conditions. They found little correspondence between the pot experiment in the greenhouse and the field experiments and concluded that extrapolation from pot experiments to normal field conditions was inappropriate. In contrast, Gove et al. (2007) tested woodland species individually in pots for 10 weeks in greenhouses and the same species individually in pots for one year under field conditions at doses equivalent to drift levels (25 % or less of label rate; see Holterman et al., 1997; de Snoo and de Wit, 1998; Weisser et al., 2002). He then compared the occurrence of the same species in a survey of 90 woodland margins, 30 each adjacent to unimproved grassland, improved grassland and arable land corresponding to low, medium and high agricultural intensity. It was found that, although there was considerable agreement between the short- and long-term single-pot experiment, the former underestimated the latter. Furthermore, in the survey of woodlot margins, the most sensitive species were lowest when abutted to high-input fields and highest near low-input fields.

Strandberg et al. (2007) compared dose–response of single-species (*Agrostis capillaris*, *Festuca ovina*) with two-species competitive interactions and later on the performance of these species in multi-species experimental grassland exposed a combination of low dosages of glyphosate (0–25 % label rate of 1 440 g a.s./ha) and nitrogen (0–100 kg N/ha). In the single-species test, *F. ovina* (ER_{10} = 35.1 g a.s./ha, ER_{50} = 114.4 g a.s./ha) was significantly less sensitive to glyphosate than *A. capillaris* (ER_{10} = 19.0 g a.i./ha, ER_{50} = 37.5 g a.s./ha). In the two-species competition study, *A. capillaris* showed little intraspecific competition, i.e. when the density of *F. ovina* = 0, and the calculated ER_{10} was comparable with the ER_{10} estimated in the single-species test. However, when *A. capillaris* grew together with *F. ovina* in a 1:1 mixture of varying densities of both species, the ER_{10} for *A. capillaris* were 16 % lower than estimated in the single-species test. This clearly demonstrates that the sensitivity of *A. capillaris* to glyphosate was affected by the presence of the less sensitive *F. ovina*.

The application of both glyphosate and nitrogen had significant effects on the vegetation at the experimental grassland (Strandberg et al., 2007). The experiment was started in 2001 and, over the years, the vegetation gradually changed with respect to both species richness and species composition. Generally, the application of glyphosate as well as nitrogen affected the species richness negatively. However, at the highest nitrogen level (100 kg N/ha), the application of low dosages of glyphosate to some extent counteracted the negative effect of nitrogen on species richness. The way glyphosate appears to compensate for the negative effect of nitrogen on species richness may be explained by glyphosate reducing the competition by the dominant species and thereby creating space for other species. The negative effect of nitrogen on species richness is well documented in the literature (e.g. Bobbink et al., 1998; Gough et al., 2000; Stevens et al., 2004; Clark and Tilman, 2008) and pesticide applications have been hypothesised as being one of the main reasons for the biodiversity decline in agricultural areas in Europe (e.g. Green, 1990; Fuller et al., 1995; Andreasen et al.,, 1996; Rich and Woodruff, 1996; Chamberlain et al., 2000; Donald et al., 2000; Atkinson et al., 2002; Benton et al., 2002).

Despite the experimental treatment, grasses dominated the vegetation of the multi-species experiment and three grasses (*F. ovina*, *A. capillaris* and *Elytrigia repens*) made up the main part of the vegetation (Strandberg et al., 2007). The treatment, however, determined the composition of the grassland community. *F. ovina* dominated the plots receiving the highest dosages of glyphosate, i.e. with application of 360 g a.i/ha, both at low-, intermediate- and high-nitrogen applications. In addition, it dominated the low-nitrogen plots with no application of glyphosate. *E. repens* dominated the vegetation in plots receiving high levels of nitrogen and low- or intermediate-glyphosate application

and *A. capillaris* did best at low and intermediate application of both glyphosate and nitrogen and it seemed to be sensitive to competition from both *F. ovina* and *E. repens*.

A comparison between dose–response single-species tests in the greenhouse and outdoors as well as the same species grown in mesocosms in the greenhouse revealed contradictory results and was either species or herbicide dependent (Riemens et al., 2008, 2009). With glufosinate ammonium (Riemens et al., 2008), plants growing individually in pots in greenhouses were more sensitive than plants growing individually in pots outdoors. However, plants grown in a mixture demonstrated much higher sensitivity than the single-species situations. With the herbicide tepraloxydim, the response of single-species tests and of the same species grown in mixture was species and dose dependent (Riemens et al., 2009). The conclusion reached from these two studies was that it is incorrect to extrapolate from single-species tests to species growing in mixtures. Likewise, plant response to herbicides growing individually in pots or within communities show different sensitivity and in many cases was shown to be species specific (Höllrigl-Rosta et al., 2005; Siemoneit-Gast et al., 2007; Reuter and Siemoneit-Gast, 2006, 2007).

Dalton and Boutin (2010) initiated two dose–response studies aimed at comparing single-species tests with the same species growing within microcosms placed both in the greenhouse and outdoors, using nine terrestrial wild species in one experiment and seven wetland species in a second experiment. In the case of terrestrial species, single-species tests were usually not the most sensitive when compared with plants growing in greenhouse microcosms, indicating that they did not represent the worst-case scenario. There was more variability with the wetland species but, in some cases, single-species tests underestimated sensitivity. When comparing greenhouse and outdoor microcosms with the same plant species, it was shown that, although outdoor plants were generally less sensitive than plants grown in the greenhouse, the latter did not capture the large variability encompassed in microcosms subjected to semi-natural outdoor conditions (Dalton and Boutin, 2010). The obvious conclusion reached was that, for an accurate risk assessment, the experimental design should be able to accommodate this natural variability (Cousens et al., 1988). In addition, changes in community structure were observed in herbicide-treated microcosms that could not be predicted from single-species testing.

7.2. Reasons for discrepancies between single- and multi-species tests

Numerous reasons could explain the discrepancy between single-species tests, frequently conducted under more controlled greenhouse conditions, and multi-species tests, often performed outdoors. Differences in sensitivity have been attributed to a multitude of confounding variables, including plant traits, phenological stage at spray, environmental factors and growing conditions.

Plant traits and anatomy have been shown to vary greatly depending on existing growth conditions. Mokany and Ash (2008) measured 14 traits on 17 species growing under greenhouse conditions (fertilised and unfertilised) and in field situations. It was found that most traits differed and concluded that pot-grown plants could not be used to quantify traits for field-grown plants. Traits such as leaf area, hairs, leaf angle and measures of growth greatly varied among growing conditions. Cuticle thickness and more active growth in the greenhouse or in the field were also seen as important variables (Clark et al., 2004). A study with *Arabidopsis thaliana* showed that total leaf area and trichome density (as well as growing conditions) played an important role in the species sensitivity to glyphosate (Boutin et al., 2012).

Field multi-species tests found in the literature, carried out for efficacy assessment and agronomic purposes, were often conducted at later phenological stages than greenhouse single-species tests, and this may be another confounding factor that is essential to consider in risk assessments. In the field, species and even individuals within a species are often at markedly different growth stages and hence may differ in susceptibility. Plants sprayed with herbicides during reproductive stages (flower buds or seed production) often exhibited more sensitivity than when sprayed at the seedling stage (Boutin et al., 2014).

Additional interacting factors such as the presence of phytophagous insects and diseases are also common under field conditions but are uncommon and controlled under greenhouse conditions. Plant density also appears to be an important aspect to take into account. Phytotoxicity seems to decrease as plant density increases and this is attributed to the amount of herbicide (or other toxins) available to each plant in a restricted volume of soil; with more plants there is less herbicide active ingredient available to each one (Hoffman and Lavy, 1978; Weidenhamer et al., 1989).

It has been suggested that, when plants are stressed or weakened or injured by toxicants, air pollution or predators, they are likely to succumb to diseases or be out-competed by more tolerant species, thus further confounding field responses (Wang and Freemark, 1995). On the other hand, when metabolism is increased and plants are growing fast, more toxicity can be observed. For example, phytotoxicity was more pronounced when plants were grown under higher nitrogen levels than with reduced fertilisation (Joyner and Couch, 1976) but responses may be species and herbicide specific (Allison et al., 2013).

Discrepancies may be related to ambient conditions during growth and at spray (Garrod, 1989). Low relative humidity and reduced light intensity were all found or assumed to influence plant sensitivity of herbicides (Mathiassen et al., 2000; Petersen and Hurle, 2001; Riemens et al., 2008, 2009). Conversely, elevated temperature (35 °C) induced higher phytotoxicity on grass species when treated with fungicides than when placed under lower temperature (22 °C) (Joyner and Couch, 1976). Boutin et al. (2010) demonstrated high variability in plant response to glyphosate and atrazine (see below).

At this point in time and with the limited studies available, it is difficult to determine if differences between single- and multi-species tests are better explained by differences in biotic or abiotic factors, and it is impossible to predict what direction the sensitivity takes. Nevertheless, multi-species tests are considered more representatives of natural conditions and have been performed with minimal effort and little difficulty, although they are more time consuming and expensive. Controlled microcosm studies may be a way to alleviate complexities related to outdoor more variable field studies (Dammgard et al., 2008; Riemens et al., 2008; Dalton and Boutin, 2010). Pfleeger et al. (2006) demonstrated that it was possible to select native plant species relevant to an area of concern using geospatial tools and taxonomic information (see also Olszyk et al., 2008).

In conclusion, much work still needs to be undertaken to better understand the difference in phytotoxicity between simplified single-species tests and more complex multi-species experiments.

7.3. Reproducibility in phytotoxicity testing

An unexplored source of uncertainty in phytotoxicity testing is the variability in the growing conditions often found when plants are growing under greenhouses as well as outdoors. It is well known that greenhouse conditions fluctuate because they are subjected to a certain extent to prevailing outdoor climate related to temperature and sunlight. This may be unavoidable and could possibly be considered appropriate given that outdoor natural conditions also fluctuate. However, the magnitude of effect that variable conditions have on species sensitivity to herbicides requires further scrutiny in the context of regulatory risk assessment. Seven different herbaceous species from five families, including four wild species and three crops, were used in addressing effects of temporal variability on plant toxicity (Boutin et al., 2010). Six to nine doses were tested separately with a formulated product of atrazine and glyphosate using six replicates for each herbicide. In all experiments, temperatures and photosynthetically active radiation (PAR) were carefully recorded. Plants were sprayed at the three- to five-leaf stage, and the aboveground biomass was harvested 28 days after spray. The ER_{25} values were calculated for every herbicide and species where applicable.

Results demonstrated that plant species exhibited variable levels of herbicide sensitivity when grown in a greenhouse at different times of the year or when grown under greenhouses or under more uniform growth chamber conditions. In numerous cases, more than one order of magnitude difference occurred in the ER_{25} values among seasons for both crops (*Solanum lycopersicon*, atrazine; *Lactuca*

sativa, atrazine and glyphosate) and non-crops (*Geum canadense*, glyphosate; *Chrysanthemum leucanthemum* and *Rudbeckia hirta*, atrazine and glyphosate). In many cases, the 10 % threshold effect (corresponding to the accepted drift level) was dependent on the time of the year. For example, *G. canadense* was more sensitive when tested with both atrazine and glyphosate in the summer than when tested in the spring. In contrast, *C. leucanthemum* was more sensitive in the autumn than in other seasons. Minimum and maximum temperatures or light intensity (PAR) measured during the course of the experiment did not explain the variability in response of the different species.

In other studies, it was shown that test conditions induced a large variability in a given species' response to herbicides. For instance, variations in temperature and water availability modified the sensitivity of *Abutilon theophrasti* to glyphosate (Zhou et al., 2007). It was found that stressed plants (through drought or flooding) were more tolerant to glyphosate; however, sensitivity increased with cooler temperatures. Contradictory results emerged from other studies on the effects of abiotic factors, such as temperature and light, on herbicide efficacy (Anderson et al., 1993; Peterson and Hurle, 2001). All these factors have to be taken into account in phytotoxicity testing even in greenhouses, where conditions are considered relatively homogeneous compared with natural environments.

In a complex experiment, Dalton and Boutin (2010) showed that plants grown in microcosms and placed outdoors were much more variable in their response to both atrazine and glyphosate than when similar microcosms were grown under more uniform greenhouse conditions. Other studies that included single-species tests in greenhouses and the outdoors yielded conflicting results because of multiple confounding and unexplained environmental factors (Kleijn and Snoeijing, 1997; Clark et al., 2004).

In conclusion, it is generally accepted that there are differences in the activities exhibited by a given pesticide in the greenhouse and in the field. Effects of the herbicide in the field can be reduced because of environmental factors (e.g. wind, temperature, rainfall conditions), plant anatomy (e.g. cuticle thickness) and physiological states of the plant (e.g. more active growth in the greenhouse) (Garrod, 1989; Clark et al., 2004; Riemens et al., 2008; Dalton amd Boutin, 2010), although Fletcher et al. (1990) reached the opposite conclusion from their literature search. There is often a poor correspondence between greenhouse and field studies (Kleijn and Snoeijing, 1997; Clark et al., 2004; Riemens et al., 2008). Nevertheless, plants used in the greenhouse are more uniform and probably represent a worst-case scenario. In the field, species and even individuals within a species often are at markedly different growth stages and, hence, differ in susceptibility; this introduces variability in the results which makes interpretation difficult. In the assessment of products, both types of data should be considered.

More studies are needed to address the calibration in phytotoxicity between lower and higher tiers, i.e. between simplified single-species tests and more complex multi-species experiments. At this point in time, it would be premature to support the inclusion of an extrapolation factor to account for the variability in plant sensitivity under different testing conditions.

7.4. Conclusions and highlights

- Single-species tests with plants growing individually in pots or as monoculture in greenhouses are performed for phytotoxicity assessment.
- Results of these tests are used to assess the effects of herbicides on plant communities growing outdoors under natural conditions where a multitude of variable conditions may be present.
- Agronomic studies conducted at the recommended label (high) rate showed mixed results between greenhouse and field trials.
- Several confounding factors may explain the differences between single- and multi-species tests conducted with sublethal doses in toxicological studies, including plant traits, phenological stage at spray, environmental factors and growing conditions.

- More studies are needed to unravel the difference in phytotoxicity between simplified single-species tests and more complex higher tier and multi-species experiments, and to address the correlation/calibration between lower and higher tier results, including the assessment of the actual level of exposure (load captured by plant surfaces or internal concentrations) reaching the plants under the different test conditions
- Multi-species tests, although more time-consuming and expensive than single-species tests, are considered more representative of natural conditions and have been (and can be) performed with minimal effort.

8. Other issues

8.1. Considering toxicity of mixtures in the risk assessment

8.1.1. Introduction

Plant protection products are normally used in combination with additives and therefore evaluations of effects of mixtures is already part of the risk assessment procedures for PPPs according to Regulation (EU) No 1107/2009.

In addition, non-target plants, like any other organisms in the environment, could also be exposed to mixtures of biologically active compounds as a consequence of simultaneous (tank mixtures) or sequential applications of PPPs on one field or on different fields within an agricultural landscape.

The use of tank mixtures is probably widespread in Europe and is used in a variety of crops (Spruijt et al., 2010; EFSA, 2012; Glass et al., 2012; Luttik et al., 2014). The possible impact of multiple exposures resulting from tank mixes is illustrated by a study by Fryday et al. (2011), which included four different crop types (arable crops, vegetable crops, orchards and soft fruit) for three years (2006–2008). It shows that applications to 66 % of the treated arable crop area contain an average of 6.15 compounds per application. For the other three crop types, approximately 50 % of the treated area is on average treated with three different compounds per application.

No specific requirement for ecotoxicological data are included in Commission Regulation (EU) No 284/2013 with regard to tank mixtures, for which specific authorisations are sought, whereas the option to request respective data or information at least on a case-by-case basis is mentioned in the toxicology section (point 7.1.8). Applying the general concept of the data requirements, a request for experimental data on toxicity to NTTPs for such tank mixtures of herbicides would clearly be a logical consequence; where such a requirement cannot be enforced (e.g. for legal reasons), modelling of mixture toxicity with approaches as described below is recommended.

Apart from few tank mixtures, for which specific authorisations are actually sought for in risk assessment, the abovementioned exposure to mixtures is not addressed in the current risk assessment procedures. However, such exposure situations might indeed compromise the general protection goal of Regulation (EU) No 1107/2009. In particular, Commission Regulation (EU) No 284/2013 does explicitly 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”.

8.1.2. Experimental determination of product/mixture toxicity

The basic data requirements supporting the authorisation of a PPP with respect to its possible effects on NTTPs are laid down in Commission Regulations (EU) No 283/2013 and No 284/2013 specifying the data requirements for active substances and products, respectively. These documents and the OECD test guidelines for seedling emergence (No 208; OECD 2006a) and for vegetative vigour (No 227; OECD 2006b) stress that the tests may be carried out using a representative PPP instead of the active substance and that a product may include one or more active ingredients.

For PPPs containing one active substance, the requirements will often be addressed by the data submitted for the active substance when the same or a similar formulation as applied for has been used in those tests. The situation is different for PPPs containing more than one active substance. Despite the existing potential for modelling mixture toxicity (see below), it is evident from existing knowledge and experience that the standard data requirements for such PPPs should not be addressed by only this type of calculations, in particular not for the toxicity to NTTPs. Hence, a formally complete dataset for such a product should comprise tests with mono-formulations for each of the active substances in the product, and tests with the product itself, where the latter would normally be used for the risk assessment. Where a different spectrum of species has been tested in the studies of the active substance toxicity and the actual product study, the relevance of such active substance-related data for the product risk assessment should always be checked. A similar situation exists where data for another formulation with the same active substance or even the same combination of active substances as the product under assessment are available, i.e. the relevance of the data needs to be checked. Modelling approaches for mixture toxicity (see below) are valuable tools evaluating the suitability of such additional data in a product risk assessment. Nevertheless it should always be kept in mind that active substance-related data are usually derived from respective mono-formulations. If modelling approaches are based on such data, it should be assured that the active substance is driving the overall toxicity of the tested mono-formulation or that the toxicity contribution of the active substance is known.

8.1.3. Modelling of mixture toxicity

In addition to and supplementing tests, effects of mixtures may be evaluated by model approaches. Two general component-driven approaches, employing the knowledge regarding the toxicity of individual mixture components, have been established in the literature, and are based either on the concept of concentration addition (CA) or independent action (IA). The principle of CA implies that individual components of the mixture contribute to mixture toxicity in proportion to their individual concentration and potency, thus acting as dilutions of one another. CA is regarded as being applicable if individual mixture components contribute to a common outcome via similar or interrelated modes of action. On the other hand, the IA approach is based on the statistical concept of independent random events and has been suggested for prediction of joint effects of dissimilarly acting components acting in a strictly independent manner. From a mechanistic point of view, however, strict independence of action "may only rarely be relevant due to converging signalling pathways and inter-linked subsystems" (Kortenkamp et al., 2009). With respect to broad, integrating population-relevant endpoints such as reproduction, and considering the paucity of information on whether or not strict independence of action may be applicable, the European scientific committees SCHER, SCCS and SCENIHR concluded that CA may be the more appropriate default model for predicting mixture effects on the population level (SCHER, SCCS, SCENIHR, 2012).

Some examples exist for which IA provided more accurate estimates of mixture toxicity than CA (Faust et al., 2003; Kortenkamp et al., 2009, 2012). These examples are confined to unicellular organisms such as algae or bacteria. However, to date, no case has been identified where IA was more accurate and at the same time more conservative than CA (Kortenkamp et al., 2012; EFSA PPR Panel, 2013b). In ecotoxicological studies comparing the performance of the two assessment concepts, CA usually yielded more conservative predictions than IA, although the differences in estimates were small (within a factor of 5; Kortenkamp et al., 2009, 2012). Thus, available scientific evidence suggests that CA may be used as a good default tool for estimating/calculating the risk of the use of mixtures, and would be expected to provide a more conservative estimate than IA (EFSA, 2012; EFSA PPR Panel, 2013b).

It is noted that both additivity concepts (CA and IA) are based on the assumption that mixture components do not interact with one another to enhance or diminish each other's toxicity. Toxicokinetic or toxicodynamic interactions may give rise either to antagonisms or synergisms, which both may be understood as deviations from expected additivity. In particular, the prospect of potential

synergisms would imply that an additivity model, e.g. CA, might in some cases be underestimating mixture effects (EFSA, 2013).

Although examples of interactions have been described in the literature, they have been considered to be relatively rare (Kortenkamp et al., 2009; EFSA et al., 2013). In cases for which a potential for interactions exists, the likelihood of occurrence of interactions is assumed to be concentration/dose dependent (less likely to occur at doses/concentrations of individual components below the lowest individual effect levels; EFSA, 2008). In a recent review involving the analyses of mammalian studies with respect to synergism, it was noted that "in a number of positive studies, the occurrence of synergy was dose-dependent and observed only at the higher doses in the study", although some studies were identified that demonstrated synergism at doses/concentrations of mixture components close to individual no observed adverse effect levels (Boobis et al., 2011; EFSA, 2013). According to the review by Kortenkamp et al. (2009), which covered both toxicological and ecotoxicological studies, examples of interactions such as synergisms appear to be largely confined to mixtures with only a few components, with deviations from additivity predictions decreasing as the complexity of the mixture increases (Kortenkamp et al., 2009; EFSA, 2012).

Principally, the identification of determinants of potential interactions may be supported by molecular mechanistic data. However, general concepts to quantitatively predict magnitudes of anticipated interactions *in silico* are currently not available (EFSA, 2013). Nevertheless, analyses of published ecotoxicological studies investigating mixture toxicity have indicated a high average predictive power of CA: for the majority of cases, deviations between CA-based predictions of EC_{50} values and observed mixture toxicity in terms of both over- and underestimations of actual toxicity were within a factor of 3 (Kortenkamp et al., 2009). In a review of pesticide mixture toxicity studies performed with aquatic organisms, actual toxicity exceeded estimates based on CA by more than a factor of 2 in only about 5 % of the cases (Belden et al., 2007). A study dealing with the prediction of aquatic toxicity of commercial pesticide mixtures, has, however, suggested that incomplete consideration of relevant mixture components, e.g. formulation additives within pesticide products, may be a relevant factor that "reduces the reliability of mixture toxicity predictions that are based solely on the active substances in the product" (Coors and Frische, 2011).

In conclusion, based on the evidence on combined toxicity of pesticides/chemicals and the risk assessment concepts available, the concept of CA is recommended by the Panel for assessing the risk of combined exposure to the active substances in a tank mix. However, the Panel acknowledges that there is uncertainty whether interactions, i.e. synergisms, might be occurring in some cases at environmental exposures, since principally these exposures cannot be considered as low if they are efficacious.

At the moment, no standardised and accepted methods are available for assessing the risk of sequential use of different pesticides.

The recovery principle has, to date, only been used in the case of a single compound or formulation (with or without repeated applications). No methods are available for assessing the recovery after multiple uses of PPPs in the cropping season. Presently, methods for risk assessment of mixtures are under development within the EFSA working group on environmental risk assessment.

In Appendix F, details of the two model concepts and a way to compare and evaluate model deviation from test data as well as details on how to consider mixture toxicity in NTTP risk assessment are provided.

8.2. Adjuvants and co-formulants (safeners, synergists, stickers)

According to Regulation (EC) No 1107/2009, interaction between the active substance, safeners, synergists and co-formulants shall be taken into account in the evaluation of PPPs.

Regulation (EC) No 1107/2009 gives the following definitions:

(a) substances or preparations which are added to a PPP to eliminate or reduce phytotoxic effects of the PPP on certain plants are referred to as "safeners";

(b) substances or preparations which, while showing no or only weak activity as referred to in paragraph 1, can give enhanced activity to the active substance(s) in a PPP and are referred to as "synergists";

(c) substances or preparations which are used or intended to be used in a PPP or adjuvant, but are neither active substances nor safeners or synergists are referred to as "co-formulants";

(d) substances or preparations which consist of co-formulants or preparations containing one or more co-formulants, in the form in which they are supplied to the user and placed on the market to be mixed by the user with a PPP and which enhance its effectiveness or other pesticidal properties, are referred to as "adjuvants".

As stated in SANCO/10329/2002 rev. 2 final, the test substance used for terrestrial plant testing should be the lead formulation (or another formulation) because formulations contain, besides the active substance, all those components and co-adjuvants required for maximising biological activity. Mostly they alter the rate of pesticide uptake. Additionally, these components may have intrinsic toxic activity. Herbicide safeners are substances used in combination with herbicides to make them "safer", that is to reduce the effect of the herbicide on crop plants, and to improve selectivity between crop plants versus weed species being targeted by the herbicide. Herbicide safeners can be used to treat crop seeds prior to planting, or they can be sprayed on plants as a mixture with the herbicide. Safeners are applied in combination with the respective herbicides to increase selectivity. They reduce the harmful effects for crop plants, whereas the phytotoxic effect on non-crop plants is unaffected.

Efficacy of herbicidal pesticides differs a lot depending on formulation properties. Consequently, the PPP used in agricultural practice (in the intended use) is more appropriate for terrestrial plant testing than the active substance as manufactured.

8.3. Metabolites and degradation products

8.3.1. 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". The approach for identification of relevant metabolites and how to address them in the risk assessment can be found, for example, in the Guidance on Aquatic Risk assessment (EFSA PPR Panel, 2013a).

8.3.2. Relevant compartments

When assessing risks to terrestrial plants, metabolites in the following media and compartments have to be considered:

Soil: data on metabolites in soil come from the environmental fate section, including information on time course of appearance and concentration level.

Plants: information on the formation of metabolites in the plants is provided by plant metabolism studies.

Depending on the design of the plant metabolism study, metabolites present in the soil and subsequently taken up by the plant *may* be covered. It is recommended that guidance is developed on when a metabolite is covered by the plant metabolism study.

8.3.3. Definition of the residue for risk assessment

In the new data requirement for active substances (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 active substance for the degradation in soil (Commission Regulation (EU) 283/2013, point 7.1.1).

In addition to the above, the requirement for the route of degradation in soil (Commission Regulation (EU) 283/2013, point 7.1.1) indicates that the study shall "be sufficient to permit the soil residue of concern to which non-target species are or may be exposed, to be defined".

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, an estimation of exposure (PEC) is necessary for each relevant compartment, as well as information on ecotoxicity.

Information on alternative information replacing experimental studies with metabolites including identification of toxophore can be found in the Aquatic Guidance document (EFSA, 2013).

8.3.4. Alternative information replacing experimental studies

The principles for assessing metabolites should, in essence, be the same as those for active substances. However, in contrast to the active substances, 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 active substances, organisms could be considered to have been exposed to the metabolites. However, this extrapolation is only valid if it is shown that the plants were exposed to a realistic or worst-case exposure profile of the metabolite. 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. Another possibility could be to prolong the test in order to lengthen the observation phase from effects occurring owing to exposure to the metabolite.

In toxicity studies with intensive lighting, it is likely that metabolites which are formed as a result of photolysis are present in an amount which is relevant for field conditions and additional toxicity testing with metabolites detected in the photolysis study might not be warranted. These conclusions should be supported by analytical measurements and the risk resulting from the metabolite can be addressed as above.

Substances that have a specific mode of action, such as chemical 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.

Detailed information on alternative information replacing ecotoxicological studies with metabolites can be found in the Aquatic Guidance document (EFSA PPR Panel, 2013a).

8.4. Addressing uncertainty in higher tier refinements and weight of evidence approaches in higher tier assessments

Examples on how to address uncertainties in risk assessment and how to use weight of evidence approaches can be found in the Risk Assessment Guidance Documents on birds and mammals (EFSA, 2009c), opinion and guidance on bees (EFSA PPR Panel, 2012b; EFSA 2013) and aquatic organisms (EFSA PPR Panel, 2013a). EFSA´s scientific committee and emerging risks unit is tasked with developing guidance on how to characterise, document and explain uncertainties in risk assessment (mandate M-2013-0261).

8.5. Research needs

Future research is needed to improve laboratory, semi-field and field tests and also to improve general knowledge on the effects of herbicides on NTTPs. The following gives details on the research needs that have been identified.

8.5.1. Species and test conditions

- Effects of herbicides on non-target terrestrial, especially perennial, species.
- Effects of herbicides on different growth stages.
- Effects of herbicides on ferns, mosses, liverworts, hornworts, horsetails, lichens and woody species.
- Measuring increasing species sensitivity in competition (intra- as well as interspecific). Establish an extrapolation factor.
- Measuring increasing species sensitivity under various abiotic conditions (light intensity, photoperiod, temperature, nutrient and organic matter levels, etc.). Establish an extrapolation factor.
- Measuring sensitivity of different crop varieties and wild plant ecotypes. Establish an extrapolation factor.

8.5.2. Effects on reproduction (flower and seed production, pollination, onset of flowering, etc.)

- Measuring endpoints other than biomass at juvenile stage.
- Effects on reproduction of plants sprayed at juvenile stage.
- Effects on reproduction of plants sprayed at flowering stage.
- Effects on F1 generation (seed germination and seedling growth).

8.5.3. Community and long-term effects

- Plant recovery after herbicide spray.
- Effects on individual species.
- Effects on populations.

- Effects on communities.
- Indirect effects of herbicide exposure on food webs, specifically on herbivores, pollinators and other flower-utilising insects, and on seed-/fruit-eating insects and birds.
- Effects of long-term repeated herbicide exposure on seed bank diversity.
- Long-term effects of recurrent herbicide applications of sublethal doses (several times a year over several years) on long-lived species (e.g. shrubs and trees).
- Effects on blooming stage (shrubs and trees).
- Effects of single and repeated herbicide exposure up the food chain on other trophic levels. Testing the validity of the risk assessment based on endpoints from the species sensitivity distribution method against field population and community data.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

- Specific protection goals have been defined in terms of spatial boundaries for off- and in-field NTTPs. SPGs should take into account ecosystem services provided by NTTPs: nutrient cycling, water regulation supporters of food webs, genetic resources and aesthetic values. The maintenance of biodiversity is also a very important SPG to be achieved. Thus, endpoints measured should be closely linked to ecosystem services provided by NTTPs. Special cases that should be considered in SPGs include NTTPs vulnerable to PPPs, conservation of rare arable weeds, protection of the NTTP seed bank, recovery from and subtle effects of sublethal doses and cascading effects at other trophic levels. Toxicity tests currently conducted for assessing pesticide effects may not be appropriate to address the SPGs described here.
- There is a paucity of data on herbicide effects on ferns, mosses, liverworts, hornworts, horsetails, lichens, mycorrhiza, terrestrial algae or woody species. For some of these groups of organisms, no information is available, and limited studies showed that they are quite sensitive and may not be sufficiently protected by current risk assessment.
- Except for the ISO methods, no standardised test guidelines for higher tier tests of herbicidal effects on NTTPs are available and the ISO methods are only applicable for testing either with contaminated soil or using a series of dilutions incorporated into control soil.
- Species selection for phytotoxicity testing of NTTP sensitivity to PPPs remains unresolved. Tests are mostly conducted with annual crop species. Research has demonstrated that they may not be adequate surrogates for wild species that need to be protected. Many non-crop species are easy to grow under greenhouse conditions and are appropriate for phytotoxicity testing. Disparities in responses between types (annuals, perennial, ferns, mosses, liverworts, hornworts, horsetails, lichens and woody species) of species, non-crop ecotypes and crop varieties are a concern and still need to be addressed. Using the plant trait-based approach is a promising avenue for plant species selection in phytotoxicity testing and ensuing ecological risk assessment. Until uncertainties surrounding species selection has been appropriately investigated, risk assessment as currently performed will remain problematic.
- Several scientific studies have been carried out and form a very useful starting point, as they include test designs that are suitable for the assessment of the importance of species interactions and test conditions. These includes (1) test designs proposed by Carpenter and Boutin (2010) and Carpenter et al. (2013) for greenhouse tests of long-term effects on reproductive endpoints of both annual and perennial species; (2) the test design for studies of herbicide effects on plant flowering (Boutin et al., 2014); and (3) general recommendations on field studies by US EPA (2012d) and recommendations for tier III field studies by Schmitz et al. (2013b). Schmitz et al. (2013b) provided a list of general recommendations on, for

example, species selection, number of species, plant densities and test duration for microcosm studies.

- The available data on both vegetative and reproductive (seed production) endpoints demonstrate the importance of considering a reproductive endpoint in regulatory assessment or use of an extrapolation factor to compensate for higher sensitivity of reproductive endpoints.
- It would be preferable to prolong vegetative growth studies to provide also the reproductive endpoints of the application of pesticides to increase the number of studies available for assessing extrapolation factors.
- Plant screening data (pre-screening, efficacy and crop margin of safety data) are very useful in risk assessment for PPPs, even though testing is not conducted under GLP.
- Drift during application is currently considered to be the most important factor for off-field emissions to non-target surfaces. Drift is normally defined as droplet drift but vapour drift can also contribute in particular cases. Exposure models to calculate loadings caused by droplet and vapour drift are presently available. However, spray drift values in field crops originating from recent research were considerably higher than those currently used in exposure assessments at the EU level.
- Dust drift is considered to be an important emission route in particular cases. However, no validated models are available so far. As dust drift of very small particles can behave in a similar way to vapour drift, it can be proposed as a starting point for dust deposition on soil (EFSA PPR Panel, 2012).
- Experiences from the exposure assessments of surface waters show that surface run-off may also contribute significantly to the contamination of non-target terrestrial ecosystems in the neighbourhood of agricultural areas. Models to estimate run-off losses are available and are used for the assessment of the aquatic environment. However, the information on vegetated buffer strips used currently in aquatic risk assessment has to be re-evaluated with regard to worst-case situations for non-target plants. The exposure via soil residues is only of relevance for seedling emergence and root uptake.
- A wide range of modelling approaches are available for linking the effects of herbicides on plant performance and population behaviour. These include models for annual plant populations and for perennials populations with size structure, such as matrix models and integral projection models. These models, however, require the estimation of growth, survival and reproduction rates in the field and so are difficult to apply when only data on growth or reproduction from short-term pot studies are available. A simple annual plant population model is presented, which demonstrates that the effects of herbicides are strongly mediated by the timing of applications relative to the action of density dependence and habitat fertility.
- A variety of risk mitigation options for in- and off-field risks are available.
- Non-target terrestrial plants are exposed at various phenological stages. Therefore, sensitivity not only is species dependent but also varies at different phenological stages: seedling, juvenile, vegetative and reproductive stages. The whole life cycle has to be considered.

RECOMMENDATIONS FOR FURTHER RESEARCH

The PPR Panel gives the following recommendations:

- There is a need for more studies on the effects of combinations of pesticides in tank mixtures and sequential applications in order to assess risk within one or more successive growing seasons.
- There is a need to develop methods to assess recovery after multiple uses of PPPs.

- The assumptions behind the derivation of the extrapolation factor (EF), especially regarding the approach of dealing with small datasets, need to be verified (see Appendix A).
- An assessment factor (AF) needs to be applied to the endpoint used to cover for other uncertainties (e.g. step from laboratory to field, single species to biocoenosis) in the risk assessment.
- Both factors (EF and AF) should be calibrated on the basis of data from the reference tier (actual ecosystem or surrogate reference tier).
- A quantitative link between the outcome of the risk assessment (TER values) and the consequences for other groups (e.g. non-target arthropods including pollinators, birds and mammals) via the food web or alteration of habitats should be derived.
- New spray drift curves should be evaluated when they become available and the spray drift assessment methodology should begin to be revised accordingly. For the time being, the PPR panel recommends the use of the current assessment based on FOCUS (2001).
- The estimation of vapour drift deposition by the EVA 2 model needs to be improved and the option to use alternative modelling approaches should be investigated.
- Conditions and methodology in the OECD guidelines (208 and 207) will need to be revisited in terms of organic matter levels, number of species per pot, duration of test, species selection and endpoints measured (including reproductive endpoints).
- Some plant groups, e.g., ferns and mosses, are sensitive to certain plant protection products, like herbicides, but the current database is insufficient to include them into a the testing and assessment framework for non-target terrestrial plants. The same consideration does apply to non-animal taxa that are important for terrestrial biodiversity, such as lychens and mushrooms. Therefore, more research is needed, including on appropriate endpoints, in order to include these plant groups as well as lychens and mushrooms into a science-based risk assessment for non-target organisms.

DOCUMENTATION PROVIDED BY EFSA

The following documents in particular are relevant to the questions raised:

1. Guidance Document on Terrestrial Ecotoxicology (SANCO/10329/2002 rev 2 final, 17 October 2002).

2. EFSA (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 – Fate and Behaviour in the Environment. 31 January 2007. The EFSA Journal 2007, 448, 1–17.

3. EFSA (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 – Ecotoxicological studies. 7 March 2007. The EFSA Journal 2007, 461, 1–44.

4. EFSA (2009). 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

5. EFSA (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

REFERENCES

Abernathy JR, Ray LL and Keeling JW, 1979. Cotton cultivar response to propazine and atrazine. Agronomy Journal, 71, 929–931.

Albrecht H, 2005. Development of arable weed seedbanks during the 6 years after the change from conventional to organic farming. Weed Research, 45, 339–350.

Albrecht H and Mattheis A, 1998. The effects of organic and integrated farming on rare arable weeds on the Forschungsverbund Agrar-ökosysteme München (FAM) research station in southern Bavaria. Biological Conservation, 86, 347–356.

Al-Khatib K and Tamhane A, 1999. Dry Pea (*Pisum sativum* L.) response to low rates of selected foliar- and soil-applied sulfonylurea and growth regulator herbicide. Weed Technology, 13, 753–758.

Allison JE, Boutin C and Carpenter D, 2013. Influence of soil organic matter on herbicide effect on wild and crop species. Ecotoxicology Oct, 22(8), 1289-302. Doi:10.1007/s10646-013-1116-3.

Alterra, PBL, RIVM (2011), http://www.pearl.pesticidemodels.eu/

Anderson DM, Swanton CJ, Hall JC and Mersey BGY, 1993. The influence of temperature and relative humidity on the efficacy of glufosinate ammonium. Weed Research, 33, 139–148.

Andreasen C and Stryhn H, 2008. Increasing weed flora in Danish arable fields and its importance for biodiversity. Weed Research, 48, 1–9.

Andreasen C, Stryhn H and Streibig JC, 1996. Decline of the flora in Danish arable fields. Journal of Applied Ecology, 33, 619–626.

Arsenault WJ and Ivany JA, 2001. Response of several potato cultivars to metribuzin and diquat. Crop Protection, 20, 547–552.

Atkinson PW, Fuller RJ and Vickery JA, 2002. Large-scale patterns of summer and winter bird distribution in relation to farmland type in England and Wales. Ecography, 25, 466–480.

Aude E, Tybirk K and Pedersen MB, 2003. Vegetation diversity of conventional and organic hedgerows in Denmark. Agriculture, Ecosystems and Environment, 99, 135–147.

Baird DJ, Rubach MN and Van den Brink PJ, 2008. Trait-based ecological risk assessment (TERA): the new frontier? Integrated Environmental Assessment and Management, 4, 2–3.

Ball DA, 1992. Weed seedbank response to tillage, herbicides, and crop rotation sequence. Weed Science, 40, 654–659.

Bedos C, Cellier P, Calvet R, Barriuso E and Gabrielle B, 2002. Mass transfer of pesticides into the atmosphere by volatilization from soils and plants: overview. Agronomie, 22, 21–33.

Belden JB, Gilliom RJ and Lydy MJ, 2007. How well can we predict the toxicity of pesticide mixtures to aquatic life? Integr Environ Assess Manag 3, 364-372.

Benton TG, Bryant DM, Cole L and Crick HQP. 2002. Linking agricultural practice to insect and bird populations: a historical study over three decades. Journal of Applied Ecology, 39, 673–687.

Benzing DH and Burt KM, 1970. Foliar permeability among twenty species of the Bromeliaceae. Bulletin of the Torrey Botanical Club, 97, 269–279.

Bernhardt-Römermann M, Gray A, Vanbergen AJ, Berges L, Bohner A, Brooker RW, De Bruyn L, De Cinti B, Dirnbock T, Grandin U, Hester AJ, Kanka R, Klotz S, Loucougaray G, Lundin L, Matteucci G, Meszaros I, Olah V, Preda E, Prevosto B, Pykala J, Schmidt W, Taylor ME,

Vadineanu A, Waldmann T and Stadler J, 2011. Functional traits and local environment predict vegetation responses to disturbance: a pan-European multi-site experiment. Journal of Ecology, 99, 777–787.

Bhatti MA, Felsot AS, Al-Khatib K, Kadir S and Parker R, 1995. Effects of simulated chlorsulfuron drift on fruit yield and quality of sweet cherries (*Prunus avium* L.). Environmental Toxicology and Chemistry, 14, 537–544.

Biesmeijer JC, Roberts SPM, Reemer M, Ohlemüller R, Edwards M. Peeters T, Schaffers AP, Potts SG, Kleukers R, Thomas CD, Settele J and Kunin WE, 2006. Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. Science, 313, 351–354.

Bilz M, Kell SP, Maxted N and Lansdown RV, 2011. European Red List of Vascular Plants. Publications Office of the European Union, Luxembourg.

Blackburn LG and Boutin C, 2003. Subtle effects of herbicide use in the context of genetically modified crops: a case study with glyphosate. Ecotoxicology, 12, 271–285.

Bleeker W, Schmitz U and Ristow M, 2007. Interspecific hybridisation between alien and native plant species in Germany and its consequences for native biodiversity. Biological Conservation, 137, 248–253.

Bobbink R, Hornung M and Roelofs JGM, 1998. The effects of air-borne nitrogen pollutants on species diversity in natural and semi-natural European vegetation. *Journal of Ecology*, **86**, 717–738.

Boobis A, Budinsky R, Collie S, Crofton K, Embry M, Felter S, Hertzberg R, Kopp D, Mihlan G, Mumtaz M, Price P, Solomon K, Teuschler L, Yang R, Zaleski R, 2011. Critical analysis of literature on low-dose synergy for use in screening chemical mixtures for risk assessment. Critical Reviews in Toxicology 41, 369-383.

Boutin C, 2006. Comparison of the vegetation and seedbanks of soybean fields, adjacent boundaries, and hedgerows in Ontario. Canadian Journal of Plant Science, 86, 557–567.

Boutin C and Jobin B, 1998. Intensity of agricultural practices and effects on adjacent habitats. Ecological Applications, 8, 544–557.

Boutin C and Rogers CA, 2000. Pattern of sensitivity of plant species to various herbicides—An analysis with two databases. Ecotoxicology, 9, 255–271.

Boutin C, Freemark KE and Keddy CJ, 1993. Proposed guidelines for registration of chemical pesticides: non-target plant testing and evaluation. Technical Report Series No. 145, Canadian Wildlife Service, Environment Canada, Gatineau, QC, 91 pp.

Boutin C, Freemark KE and Keddy CJ, 1995. Overview and rationale for developing regulatory guidelines for nontarget plant testing with chemical pesticides. Environmental Toxicology and Chemistry, 14, 1465–1475.Boutin C, Lee H-B, Peart ET, Batchelor SP and Maguire RJ, 2000. Effects of the sulfonylurea herbicide metsulfuron methyl on growth and reproduction of five wetland and terrestrial plant species. Environmental Toxicology and Chemistry, 19, 2537–2541.

Boutin C, Elmegaard N and Kjær C, 2004. Toxicity testing of fifteen non-crop plant species with six herbicides in a greenhouse experiment: implications for risk assessment. Ecotoxicology, 13, 349–369.

Boutin C, White AL and Carpenter D, 2010. Measuring variability in phytotoxicity testing using crop and wild plant species. Environmental Toxicology and Chemistry, 29, 237–242.

Boutin C, Baril A, McCabe SK, Martin PA and Guy M, 2011. The value of woody hedgerows for moth diversity on organic and conventional farms. Environmental Entomology, 40, 560–569.

Boutin C, Aya KL, Carpenter D, Thomas PJ and Rowland O, 2012. Phytotoxicity testing for herbicide regulation: shortcomings in relation to biodiversity and ecosystem services in agrarian systems. Science of the Total Environment, 415, 79–92.Boutin C, Strandberg B, Carpenter D, Mathiassen

SK and Thomas P, 2014. Herbicide impact on non-target plant reproduction: what are the toxicological and ecological implications? Environmental Pollution, 185, 295–306.

Boutin C, Strandberg B, Carpenter D, Mathiassen SK, Thomas P, 2014. Herbicide impact on native plant reproduction: what are the ecological and toxicological implications? Environmental Pollution. 185: 1-12.

Breeze VG, Marshall EJP, Hart A, Vickery JA, Crocker J, Walters K, Packer J, Kendall D, Fowbert J and Hodkinson D, 1999. Assessing pesticide risks to non-target terrestrial plants. Pesticides Safety Directorate, Commission No PN0923. Ministry of Agriculture, Fisheries and Food, London, UK.

Brewer, CE and Oliver LR, 2007. Weed seed rain with late-season glyphosate applications. Weed Technology, 21, 733–758.

Bright JA, Morris AJ and Winspear R, 2008. A review of indirect effects of pesticides on birds and mitigating land-management practices. Royal Society for the Protection of Birds, Sandy, UK, 1–66.

Brown RA and Farmer D, 1991. Track-sprayer and glasshouse techniques for terrestrial plant bioassays with pesticides. In Plants for toxicity assessment, 2nd volume. STP 1115. Eds Gorsuch JW, Lower WR, Wang W and Lewis MA. American Society for Testing and Materials, West Conshohocken, PA, 197–208.

Burton JD, Monks DW, Maness EP and Robinson DK, 1994. Sulfonylurea selectivity and safener activity in 'Landmark' and 'Merit' sweet maize. Pesticide Biochemistry and Physiology, 48, 172.

Campbell LH and Cooke AS, 1997. The indirect effects of pesticides on birds. Joint Nature Conservation Committee, Peterborough, UK, 18 pp.

Caswell H, 2001. Matrix population models. construction, analysis and interpretation. Sinauer Associates, Sunderland, MA.

Carpenter D and Boutin C, 2010. Sublethal effects of the herbicide glufosinate ammonium on crops and wild plants: short-term effects compared to vegetative recovery and plant reproduction. Ecotoxicology, 19, 1322–1336.

Carpenter D, Boutin C and Allison JE, 2013. Unforeseen effects of the herbicide chlorimuron ethyl on wetland and terrestrial plants: time to vegetative and reproductive recovery following exposure. Environmental Pollution, 172, 275–282.

Cathcart RJ, Chandler K and Swanton CJ, 2004. Fertilizer nitrogen rate and the response of weeds to herbicides. Weed Science, 52, 291–296.

Cavers PB, 1995. Seed banks: memory in soil. Canadian Journal of Soil Science, 75, 11–13.

Chachalis D, Reddy KN, Elmore CD and Steele ML, 2001a. Herbicide efficacy, leaf structure, and spray droplet contact angle among *Ipomoea* species and smallflower morninglory. Weed Science, 49, 628–634.

Chachalis D, Reddy KN and Elmore CD. 2001b. Characterization of leaf surface, wax composition, and control of redvine and trumpetcreeper with glyphosate. Weed Science, 49, 156–163.

Chamberlain DE, Fuller RJ, Bunce RGH, Duckworth JC and Shrubb M, 2000. Changes in abundance of farmland birds in relation to the timing of agricultural intensification in England and Wales. Journal of Applied Ecology, 37, 771–788.

Clark CM and Tilman D, 2008. Lost of plant species after chronic low-level nitrogen deposition to prairie grassland. Nature, 451, 712–715.

Clark J, Ortego LS and Fairbrother A, 2004. Sources of variability in plant toxicity testing. Chemosphere, 57, 1599–1612.

Clay PA and Griffin JL, 2000. Weed seed production and seedling emergence responses to late-season glyphosate applications. Weed Science, 48, 481–486.

Cognard J, 2013. Etude de l'impact du glyphosate sur *Arabidopsis thaliana*. Rapport de stage. Agrocampus Ouest, Centre de Rennes, France, 39 pp.

Cole JFH, Canning L and Brown RA, 1993. Rationale for the choice of species in the regulatory testing of the effects of pesticides on terrestrial non-target plants. Proceedings Brighton Crop Protection Conference—Weeds, Brighton, UK, 22–25 November, 151–156.

Comas LH and Eissenstat DM, 2009. Patterns in root trait variation among 25 co-existing North American forest species. Phytologist, 182, 919–928.

Convention for Biological Diversity, 2010. Global Strategy for Plant Conservation. Available online: www.bgci.org/resources/news/0753/

Coors A and Frische T, 2011. Predicting the aquatic toxicity of commercial pesticide mixtures. Environmental Sciences Europe 23, 22

Cornish PS and Burgin S, 2005. Residual effects of glyphosate herbicide in ecological restoration. Restoration Ecology, 13, 695–702.

Cousens R, 1991. Aspects of the design and interpretation of competition (interference) experiments. Weed Technology, 53, 664–673.

Cousens R, Marshall EJP and Arnold GM, 1988. Problems in the interpretation of effects of herbicides on plant communities. In British Crop Protection Council, Environmental Effects of Pesticides. BCPC Monograph No. 40.

Critchley CNR, Fowbert JA and Sherwood AJ, 2006. The effects of annual cultivation on plant community composition of uncropped field boundary strips. Agriculture, Ecosystems and Environment, 113, 196–205.

Crone EE, Marler M and Pearson DE, 2009. Non-target effects of broadleaf herbicide on a native perennial forb: a demographic framework for assessing and minimizing impacts. Journal of Applied Ecology, 46, 673–682.

Dalton RL and Boutin C, 2010. Evaluation of phytotoxicity testing: comparing the effects of herbicides on non-target plants grown singly and in microcosms. Environmental Toxicology and Chemistry, 29, 2304–2315.

Damgaard C, 2003. Modelling plant competition along an environmental gradient. Ecological Modelling, 170, 45–53.

Damgaard C, 2011. Measuring competition in plant communities where it is difficult to distinguish individual plants. Computational Ecology and Software, 1, 125–137.

Damgaard C, Mathiassen SK and Kudsk P, 2008. Modelling effects of herbicide drift on the competitive interactions between weeds. Environmental Toxicology and Chemistry, 27, 1302–1308.

Damgaard C, Strandberg, B, Strandberg MT, Aude E, Sørensen PB, Nielsen KE, Bruus M, 2014. Selection on plant traits in hedgerow ground vegetation: the effect of time since conversion from conventional to organic farming. Basic and Applied Ecology, 15, 250-259.

DEFRA (Department for the Environment, Food and Rural Affairs), 2004. Development of a scheme for the assessment of risks to wider biodiversity arising from the use of pesticides. DEFRA Research Project PS2403. Final report. Available online: http://randd.defra.gov.uk

DEFRA (Department for the Environment, Food and Rural Affairs), 2005. Assessing the indirect effects of pesticides on birds. Final report PN0925. http://randd.defra.gov.uk/Document.aspx?Document=PN0925_2486_FRP.pdf

de Jong FMW and de Haes HAU, 2001. Development of a field bioassay for the side-effects of herbicides on vascular plants using *Brassica napus* and *Poa annua*. Environmental Toxicology, 16, 397–407.

de Lange HJ, Lahr J, Van der Pol JJC, Wessels Y and Faber JH. 2009. Ecological vulnerability in wildlife. An expert judgment and multi-criteria analysis tool using ecological traits to assess relative impact of pollutants. Environmental Toxicology and Chemistry, 28, 2233–2240.

de Snoo GR, 1997. Arable flora in sprayed and unsprayed crop edges. Agriculture, Ecosystems and Environment, 66, 223–230.

de Snoo GR and de Wit PJ, 1998. Buffer zones for reducing pesticide drift to ditches and risks to aquatic organisms. Ecotoxicology and Environmental Safety, 41, 112–118.

de Snoo GR, Tamis WLM and van de Brink PJ, 2005. Non-target plant field study: effects of glufosinate ammonium on off crop vegetation. Institute of Environmental Sciences (CML), Report 161, Leiden University, Netherlands.

de Weese WW, Wax LM, Carlson WC and Ciarletta JA, 1989. Response of soybean (*Glycine max*) cultivars to metribuzin in the field and greenhouse. Weed Technology, 3, 566–572.

DeGennaro F and Weller SC, 1984. Differential susceptibility of field bindweed (*Convolvulus arvensis*) biotypes to glyphosate. Weed Science, 32, 472–476.

Díaz S and Cabido M, 2001. Vive la différence: plant functional diversity matters to ecosystem processes. Trends in Ecology and Evolution, 16, 645–655.

Donald PF, Green RE and Heath MF, 2000. Agricultural intensification and the collapse of Europe's farmland bird populations. Proceedings of the Royal Society London Series B, 268, 25–29. doi: 0.1098/rspb.2000.1325.

Dorrough J and Scroggie MP, 2008. Plant responses to agricultural intensification. Journal of Applied Ecology, 45, 1274–1283.

Easterling MR, Ellner SP and Dixon PM, 2000. Size-specific sensitivity: applying a new structured population model. Ecology, 81, 694–708.

EC (European Commission), 2002a. Guidance Document on Terrestrial Ecotoxicology in the context of the Directive 91/414/EEC. SANCO/10329/2002, rev.2 final, 17 October 2002, pp. 1–39.

EC (European Commission), 2002b. Guidance Document on Risk Assessment for Birds and Mammals under Council Directive 91/414/EEC. SANCO/4145/2000, final 25 September 2002, 74.

EC (European Commission), 2014. Draft authorisation of Plant Protection Products for Seed Treatment (draft), SANCO/10553/2012, January 2014

EFSA (European Food Safety Authority), 2004. Opinion of the Scientific Panel on Plant Health, Plant Protection Products and their Residues on a request of EFSA related to FOCUS SW scenarios. The EFSA Journal 2004, 145, 1–31.

EFSA (European Food Safety Authority), 2007. Opinion of the Scientific Panel on Plant protection products and their Residues on the Final Report of the FOCUS Air Working Group on Pesticides in Air: Consideration for exposure assessment. (SANCO/10553/2006 draft 1 (13 July 2006))The EFSA Journal 2007, 513, 1-30.

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, 1–181. doi: 10.2903/j.efsa.2008.734.

EFSA (European Food Safety Authority), 2009. 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), 2013. Guidance on the risk assessment of plant protection products on bees (*Apis mellifera*, *Bombus* spp. and solitary bees). EFSA Journal 2013;11(7):3295, 266 pp. doi: 10.2903/j.efsa.2013.3295.

EFSA (European Food Safety Authority), 2014. Guidance on the estimation of DegT50 values. EFSA Journal 2014;12(5):3662, 37 pp.

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.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2012a. 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 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. doi:10.2903/j.efsa.2012.2668.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2013a. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. EFSA Journal 2013;11(7):3290, 268 pp. doi: 10.2903/j.efsa.2013.3290.

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2013b. Scientific Opinion on relevance of dissimilar mode of action and its appropriate application for cumulative risk assessment of pesticides residues in food. EFSA Journal 2013;11(12): 3472, 40 pp. doi:10.2903/j.efsa.2013.3472

Eggers T, 1987. Environmental impact of chemical weed control in arable fields in the Federal Republic of Germany. Proceedings 1987 Brighton Crop Protection Conference—Weeds, Brighton, UK, 267–275.

Ehler, LE, 2006 ; Integrated pest management (IPM): definition, historical development and implementation, and the other IPM; Pest Manag Sci 62:787–789.

Ellner SP and Rees M, 2006. Integral projection models for species with complex demography. American Naturalist, 167, 410–428.

EPPO Guideline PP 1/207, 2007. Efficacy evaluation of plant protection products. Effects on succeeding crops, Bulletin OEPP/EPPOBulletin 37, 452–498.

Faust M, Altenburger R, Backhaus T, Blanck H, Bödeker W, Gramatica P, Hamer V, Scholze M, Vighi M and Grimme LH, 2003. Joint algal toxicity of 16 dissimilarly acting chemicals is predictable by the concept of independent action. Aquatic Toxicology 63, 43-63.

Fischer A and Milberg P, 1997. Effects on the flora of extensified use of field margins. Swedish Journal of Agricultural Research, 27, 105–111.

Fletcher JS, Muhitch MJ, Vann DR, McFarlane JC and Benenati FE, 1985. Phytotox database evaluation of surrogate plant species recommended by the U.S. environmental protection agency and the Organisation for Economic Co-operation and Development. Environmental Toxicology and Chemistry, 4, 523–532.

Fletcher JS, Pfleeger TG and Ratsch HC, 1993. Potential environmental risks associated with the new sulfonylurea herbicides. Environmental Science and Technology, 27, 2250–2252.

Fletcher JS, Johnson FL and McFarlane JC, 1990. Influence of greenhouse versus field testing and taxonomic differences on plant sensitivity to chemical treatment. Environmental Toxicology and Chemistry, 9, 769–776.

Fletcher JS, Pfleeger TG and Ratsch HC, 1995. Chlorsulfuron influence on garden pea reproduction. Physiologia Plantarum, 94, 261–267.

Fletcher JS, Pfleeger TG, Ratsch HC and Hayes R, 1996. Potential impact of low levels of chlorosulfuron and other herbicides on growth and yield of nontarget plants. Environmental Toxicology and Chemistry, 15, 1189–1196.

FOCUS (Forum for the Co-ordination of Pesticide Fate Models and their Use), 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 (Forum for the Co-ordination of Pesticide Fate Models and their Use), 2007. Final Report of the FOCUS Working Group on Landscape and Mitigation Factors in Ecological Risk Assessment. SANCO/10422/2005, version 2.0, September 2007.

FOCUS (Forum for the Co-ordination of Pesticide Fate Models and their Use), 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.

Frans RE and Talbert RE, 1977. Design of field experiments and the measurement and analysis of plant response. In Research methods in weed science. Ed. Truelove B. Southern Weed Science Policy, Auburn, AL, 15–23.

Fraser LH and Keddy P, 1997. The role of experimental microcosms in ecological research. Trends in Ecology and Evolution, 12, 478–481.

Freisen GH and Wall DA, 1984. Response of potato (*Solanum tuberosum*) cultivars to metribuzin. Weed Science, 32, 444.Friesen HA, Vanden Born WH, Keys CH, Dryden RD, Molberg ES and Siemens B, 1968. Effect of time of application and dosage of dicamba on tolerance of wheat, oats and barley. Canadian Journal of Plant Science, 48, 213–215.

Friesen HA, O'Sullivan PA and Vanden Born WH, 1976. HOE 23408, a new selective herbicide for wild oats and green foxtail in wheat and barley. Canadian Journal of Plant Science, 56, 567–578.

Fried G, Petit S, Dessaint F and Reboud X, 2009. Arable weed decline in Northern France: crop edges as refugia for weed conservation? Biological Conservation, 142, 238–243.

Fryday S, Thompson H and Garthwaite D, 2011. Background information for considering risk of exposure to multiple pesticides. DEFRA Project PS2354, pp 102. http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=16736

Fuller RJ, Gregory RD, Gibbons DW, Marchant JH, Wilson JD and Carter N, 1995. Population declines and range contractions among lowland farmland birds in Britain. Conservation Biology, 9, 1425–1441.

Garrod JP, 1989. Comparative responses of laboratory and field grown test plants to herbicides. Aspects in Applied Biology, 21, 51–62.

Gauvrit C and Chauvel B, 2010. Sensitivity of *Ambrosia artemisiifolia* to glufosinate and glyphosate at various developmental stages. Weed Research, 50, 503–510.

Glass R, et al., 2012. Collection and assessment of data relevant for non-dietary cumulative exposure to pesticides and proposal for conceptual approaches for non-dietary cumulative exposure assessment. Supporting Publications 2012, EN-346. Available online: http://www.efsa.europa.eu/de/supporting/doc/346e.pdf

Glover DG and Schapaugh WT Jr, 1997. Screening of soybean for pendimethalin herbicide induced stem damage. Crop Science, 37, 358.

Gotelli NJ and McCabe DJ, 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. Ecology, 83, 2091–2096.

Gough L, Osenberg CW, Gross KL and Collins SL, 2000. Fertilization effects on species density and primary productivity in herbaceous plant communities. *Oikos,* **89**, 428–439.

Gove B, Power SA, Buckley GP and Ghazoul J, 2007. Effects of herbicide spray drift and fertilizer overspread on selected species of woodland ground flora: comparison between short-term and long-term impact assessments and field surveys. Journal of Applied Ecology, 44, 374–384.

Green BH, 1990. Agricultural intensification and the loss of habitat, species and amenity in British grasslands: a review of historical change and assessment of future prospects. Grass and Forage Science, 45, 365–372.

Grime JP, 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. Journal of Ecology, 86, 902–910.

Grime JP, 2001. Plant strategies, vegetation processes and ecosystem properties. Wiley, Chichester, UK.

Gulden RH, Lewis DW, Froese JC, Van Acker RC, Martens GB, Entz MH, Derksen DA and Bell LW, 2011. The effect of rotation and in-crop weed management on the germinable weed seedbank after 10 years. Weed Science, 59, 553–561.

Guo SL, Jiang HW, Fang F and Chen, GQ, 2009. Influences of herbicides, uprooting and use as cut flowers on sexual reproduction of *Solidago canadensis*. Weed Research, 49, 291–299.

Harrison HF Jr and Keinath AP, 2003. Glasshouse assessment of clomazone response of US pumpkin cultivars. Crop Protection, 22, 795–798.

Hess FD, Bayer DE and Falk RH, 1974. Herbicide dispersal patterns: I. As a function of leaf surface. Weed Science, 22, 394–401.

Hoffman DW and Lavy TL, 1978. Plant competition for atrazine. Weed Science 26, 94–99.

Höllrigl-Rosta A, Schulte C, Reuter S, Siemoneit S and Habekost M, 2005. Protecting the environment at the edge of the field—risk assessment strategies for terrestrial non-target plants in the EU. Proceedings of the BCPC International Congress—Crop Science and Technology. British Crop Protection Council, Alton, UK, 360–370.

Höllrigl-Rosta A, Swarowsky K, Craig P, Schweikert C, 2014. Estimating the possible cumulation of effects in a risk assessment for non-target terrestrial plants (NTTPs) based on toxicity data for multiple species. Poster presentation at the SETAC Europe 24th Annual Meeting, 11-15 May 2014, Basel, Switzerland.

Holzschuh A, Steffan-Dewenter I, Kleijn D and Tscharntke T, 2007. Diversity of flower-visiting bees in cereal fields: effects of farming system, landscape composition and regional context. Journal of Applied Ecology, 44, 41–49.

Hooper DU, Chapin III FS, Ewel JJ, Hector A, Inchausti P, Lavorel S, Lawton JH, Lodge DM, Loreau M, Naeem S, Schmid B, Setälä H, Symstad AJ, Vandermeer J and Wardle DA, 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecology Monograph, 75, 3–35.

Hopen HJ, Hughes RL and Michaelis BA, 1993. Selectivity among cabbage (*Brassica oleracea* L.) cultivars by clomazone. Weed Technology, 7, 471–477.

Holterman HJ, van de Zande JC, Porskamp HAJ and Huijsmans JFM, 1997. Modelling spray drift from boom sprayers. Computers and Electronics in Agriculture, 19, 1–22.

Huangfu CH, Song XL and Qiang S. 2009. Morphological disparities in the epidermal and anatomical features of the leaf among wild *Brassica juncea* populations. Weed Biology and Management, 9, 234–242.

Huijsmans JFM and van de Zande JC, 2011. Workshop on harmonisation of drift and drift reducing methodologies for evaluation and authorization of plant protection products, Wageningen, Netherlands, 1–2 December 2010. Plant Research International, part of Wageningen UR, Business Unit Agrosystems, May 2011, Report 390.

Hull HM, Davis DG and Stolzenberg GE, 1982. Actions of adjuvant on plant surface. In: Adjuvants for herbicides. Ed. Hodgson RH, Weed Science Society of America, Champaign, Illinois, 26–67.

Hull SC Jr, 1971. Effect of spraying with 2,4-D upon abundance of pocket gophers in Franklin Basin, Idaho, Journal of Range Management, 23, 230–232.

Hulting AG, Wax LM, Nelson RL and Simmons FW, 2001. Soybean (*Glycine max* L. Merr.) cultivar tolerance to sulfentrazone. Crop Protection, 20, 679–683.

Inouye BD, 2001. Response surface experimental designs for investigating interspecific competition. Ecology, 82, 2696–2706.

Isaac MA, Murdock EC, Toler JE and Wallace SU, 1989. Effects of late-season herbicide applications on sicklepod (*Cassia obtusifolia*) seed production and viability. Weed Science, 37, 761–765.

ISO 11268-2. Soil quality—Effects of pollutants on earthworms (*Eisenia fetida*). Part 2: Determination of effects on reproduction.

ISO 11269-2. Soil quality—Determination of the effects of pollutants on soil flora. Part 2: Determination of effects on reproduction.

Jacques GI and Harvey RG, 1979. Dinitroaniline herbicides in soil. Weed Science, 27, 660–665.

Jahn T, Hötker H, Oppermann R, Bleil R and Vele L, 2014. PProtection of biodiversity of free living birds and mammals in respect of the effects of pesticides , Umweltbundesamt; Texte 30/2014; Available online: http://www.umweltbundesamt.de/publikationen/protection-of-biodiversity-of-free-living-birds

Jonason D, Andersson GKS, Ockinger E, Rundlof M, Smith HG, Bengtsson J, 2011. Assessing the effect of the time since transition to organic farming on plants and butterflies. Journal of Applied Ecology 48, 543-550

José-María L and Sans FX 2011. Weed seedbanks in arable fields: effects of management practices and surrounding landscape. Weed Research, 51, 631–640.

José-María L, Armengot L, Blanco-Moreno JM, Bassa M and Sans FX, 2010. Effects of agricultural intensification on plant diversity in Mediterranean dryland cereal fields. Journal of Applied Ecology, 47, 832–840.

Joyner BG and Couch HB, 1976. Relation of dosage rates, nutrition, air temperature, and susceptible genotype to side effects of systemic fungicides on turfgrass. Phytopathology, 66, 806–810.

Juuti S, Norokorpi Y, Helle T and Ruuskanen J, 1996. Trichloroacetic acid in conifer needles and arboreal lichens in forest environments. Science of the Total Environment, 180, 117–124.

Kahmen S and Poschlod P, 2008. Effects of grassland management on plant functional trait composition. Agriculture, Ecosystems and Environment, 128, 137–145.

Kalsch W, Junker T and Römbke J, 2006. A chronic plant test for the assessment of contaminated soils Part 1: Method development. Journal of Soil Sediment, 6, 37–45.

Kegode GO and Fronning BE, 2005. *Artemisia biennis* (biennial wormwood) control is influenced by plant size and weed flora at time of herbicide application. Crop Protection, 24, 915–920.

Keifer DW, 1989. Tolerance of corn (*Zea mays*) lines to clomazone. Weed Science, 37, 622–628.

Kjær C, 1994. Sublethal effects of chlorsulfuron on black bindweed (*Polygonum convolvulus* L.). Weed Research, 34, 453–459.

Kjær C, Strandberg M and Erlandsen M, 2006a. Metsulfuron spray drift reduces fruit yield of hawthorn (*Crataegus monogyna* L.). Science of the Total Environment, 356, 228–234.

Kjær C, Strandberg M and Erlandsen M, 2006b. Effects on hawthorn the year after simulated spray drift. Chemosphere, 63, 853–859.

Kjaer C, Bruus M, Bossi R., Lofstrom P, Andersen HV, Nuyttens D and Larsen ES, 2014. Pesticide drift deposition in hedgerows from multiple spray swaths. *Journal of Pesticide Science*, *39*(1), 14-21.

Kleijn D and Snoeijing GIJ, 1997. Field boundary vegetation and the effects of agrochemical drift: botanical change caused by low levels of herbicide and fertilizer. Journal of Applied Ecology, 34, 1413–1425.

Kleijn D and van der Voort LAC, 1997. Conservation headlands for rare arable weeds: the effects of fertiliser application and light penetration for plant growth. Biological Conservation, 81, 57–67.

Kleijn D and Verbeek M, 2000. Factors affecting the species composition of arable field boundary vegetation. Journal of Applied Ecology, 37, 256–266.

Klingaman TE and Oliver LR, 1996. Existence of ecotypes among populations of entire leaf morning glory (*Ipomoea hederacea* var. integriuscula). Weed Science, 44, 540–544.

Koch H, Weisser P and Strub O, 2004. Comparison of dose–response of pesticide spray deposits versus drift deposits. Nachrichten des Deutschen Pflanzenschutzdienstes, 56, 30–34.

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. Available online:
http://ec.europa.eu/environment/chemicals/effects/effects_en.htm

Kortenkamp A, Evans R, Faust M, Kalberlah F, Scholze M and Schuhmacher-Wolz U, 2012. Investigation of the state of the science on combined actions of chemicals in food through dissimilar modes of action and proposal for science-based approach for performing related cumulative risk assessment. Supporting Publications 2012: EN-232 [pp.]. Available online: http://www.efsa.europa.eu/de/search/doc/232e.pdf

Lagerlöf J, Stark J and Svensson B, 1992. Margins of agricultural fields as habitats for pollinating insects. Agriculture, Ecosystems and Environment, 40, 117–124.

Le Duc MG, Pakeman RJ and Marrs RH, 2000. Vegetation development on upland and marginal land treated with herbicide, for bracken (*Pteridium aquilinum*) control, in Great Britain. Journal of Environmental Management, 58, 147–160.

Leck MA, Parker VT and Simpson RL (eds), 1989. Ecology of soil seed banks. Academic Press, San Diego, CA.

Lingenfelter DD and Curran WS, 2007. Effect of glyphosate and several ACCase-inhibitor herbicides on wirestem muhly (*Muhlenbergia frondosa*) control. Weed Technology, 21, 732–738.

Longley M and Sotherton NW, 1997. Factors determining the effects of pesticides upon butterflies inhabiting arable farmland. Agriculture, Ecosystems and Environment, 61, 1–12.

Luttik R, Zorn M, Van der Brink PJ, Brock TCM, Roex E, Van der Linden AMA. 2014. Multiple stress by repeated use of plant protection products in agricultural areas. Report 601720001. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.

McCalla JH Jr, Richardson MD, Karcher DE and Boyd JW, 2004. Tolerance of seedling Bermudagrass to postemergence herbicides. Crop Science, 44, 1330–1336.

McKelvey RA, Wright JP and Honegger JL, 2002. A comparison of crop and non-crop plants as sensitive indicator species for regulatory testing. Pesticide Management Science, 58, 1174.

Marrs RH and Frost AJ, 1997. A microcosm approach to the detection of the effects of herbicide spray drift in plant communities. Journal of Environmental Management, 50, 369–388.

Marrs RH, Williams CT, Frost AJ and Plant RA, 1989. Assessment of the effects of herbicide spray drift on a range of plant species of conservation interest. Environmental Pollution, 59, 71–86.

Marrs RH, Frost AJ and Plant RA, 1991a. Effects of herbicide spray drift on selected species of nature conservation interest: the effects of plant age and surrounding vegetation. Environmental Pollution, 69, 223–235.

Marrs RH, Frost AJ and Plant RA, 1991b. Effect of mecoprop drift on some plant-species of conservation interest when grown in standardized mixtures in microcosms. Environmental Pollution, 73, 25–42.

Marrs RH, Frost AJ, Plant RA and Lunnis P, 1992. Aerial applications of asulam: a bioassay technique for assessing buffer zones to protect sensitive sites in upland Britain. Biological Conservation, 59, 19–23.

Marrs RH, Frost AJ, Plant RA and Lunnis P, 1993. Determination of buffer zones to protect seedlings of non-target plants from the effects of glyphosate spray drift. Agriculture, Ecosystems and Environment, 45, 283–293.

Marshall EJP, 1989a. Susceptibility of four hedgerow shrubs to a range of herbicides and plant growth regulators. Annals of Applied Biology, 115, 469–479.

Marshall EJP, 1989b. Distribution patterns of plants associated with arable field edges. Journal of Applied Ecology, 26, 247–258.

Marshall EJP, 2009. The impact of landscape structure and sown grass margin strips on weed assemblages in arable crops and their boundaries. Weed Research, 49, 107–115.

Marshall J, Brown V, Boatman N, Lutman P and Squire G, 2001. The impact of herbicides on weed abundance and biodiversity In: A report for the UK Pesticides Safety Directorate—PN0940: 147. British Health and Safety Executive, Chemicals Regulation Directorate Pesticides, London, UK.

Mathiassen SK, Kudsk P and Jensen PK, 2000. Comparison of herbicide performance in climate simulator, semi-field and field experiments. Proceedings of a Symposium organised by the SCI Crop Protection Group and the British Crop Protection Council Surrey, UK, University of Kent, Canterbury 10-12 Sep. 2000. Symposium Proceedings No 74 pp. 209-214.

Mayeux HS and Jordan WR, 1980. Epicuticular wax content and herbicide penetration of goldenweed leaves. Proceedings of the Weed Science Society, 33, 289.

Mayo CM, Horak MJ, Peterson DE and Boyer JE, 1995. Differential control of four Amaranthus species by six postemergence herbicides in soybean (*Glycine max*). Weed Technology, 9, 141–147.

Millennium Ecosystem Assessment (MEA), 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, DC, 160 pp.

Mokany K and Ash J, 2008. Are traits measured on pot grown plants representative of those in natural communities? Journal of Vegetation Science, 19, 119–126.

Newman MC, Ownby DR, Mezib LCA, Powell DC, Christensen TRL, Lerberg SB and Anderson B-A, 2000. Applying species-sensitivity distributions in ecological risk assessment: assumptions of distribution type and sufficient numbers of species. Environmental Toxicology and Chemistry, 19, 508–515.

Newmaster SJ and Bell FW, 2002. The effects of silvicultural disturbances on cryptogam diversity in the boreal-mixedwood forest. Canadian Journal of Forestry Research, 32, 38–51.

Noldin JA, Chandler JM, Ketchersid ML and McCauley GN, 1999. Red rice (*Oryza sativa*) biology. II. Ecotype sensitivity to herbicides. Weed Technology, 13, 19–24.

Norberg-King TJ, 1993. A linear interpolation method for sublethal toxicity: the inhibition concentration (ICp) approach (version 2.0). Technical Report 03-93. US Environmental Protection Agency, Environmental Research Laboratory, Duluth, MN.

Obrigawitch TT, Cook G, and Wetherington J, 1998. Assessment of effects on non-target plants from sulfonylurea herbicides using field approaches. Pesticide Science, 52, 199–217.

OECD (Organisation for Economic Co-operation and Development), 1984. Guidelines for testing chemicals, 2nd addendum. OECD, Paris, France.

OECD (Organisation for Economic Co-operation and Development), 2006a. Terrestrial plants test: seedling emergence and seedling growth test (No. 208). OECD Guidelines for Testing Chemicals. OECD, Paris, France.

OECD (Organisation for Economic Co-operation and Development), 2006b. Terrestrial plants test: seedling growth test (No. 208) and vegetative vigour test (No. 227). OECD Guidelines for Testing Chemicals. OECD, Paris, France.

Olszyk D, Burdick CA, Pfleeger T, Lee EH and Watrud LS, 2004. Assessing risks to non-target terrestrial plants from herbicides. Journal of Advances in Agriculture.

Olszyk D, Blakeley-Smith M, Pfleeger T, Lee EH, King G, Plocher M and Kern J, 2006. Use of vegetative vigor test to detect effects of low levels of herbicides on selected non-crop plants. Research Report. EPA/600/R-06/117. Environmental Protections Agency, Washington, DC.

Olszyk D, Pfleeger T, Lee EH, Burdick C, King G, Plocher M and Kern J, 2008. Selecting and evaluating native plants for region-specific phytotoxicity testing. Integrated Environmental Assessment and Management, 4, 105–117.

Olszyk D, Pfleeger T, Lee EH and Plocher M, 2009. Pea (*Pisium sativum*) seed production as an assay for reproductive effects due to herbicides. Environmental Toxicology and Chemistry, 28, 1920–1929.

Pallett K, Cole J, Oberwalder C and Porch J, 2007. Performance of potential non-crop or wild species under OECD 208 testing guideline study conditions for terrestrial non-target plants. Pest Management Science, 63, 134–140.

Perry NH, Chaney K and Wilcox A, 1996. The effect of fertiliser and herbicide application on herbaceous field margin communities. Aspects of Applied Biology, 44, 339–344.

Pestemer W and Zwerger P, 1999. Application of a standardized bioassay to estimate the phytotoxic effects of frequently used herbicides on non-target plants. XI. Symposium Pesticide Chemistry—Human and Environmental Exposure to Xenobiotics, Cremona, Italy. 763–770.

Petanidou T, Van Laere A, Ellis WN and Smets E, 2006. What shapes amino acid and sugar composition in Mediterranean floral nectars? Oikos, 115, 155–169.

Petersen J and Hurle K, 2001. Influence of climatic conditions and plant physiology on glufosinate ammonium efficacy. Weed Research, 41, 31–39.

Petersen S., Axelsen JA, Tybirk K, Aude E, Vestergaard P, 2006. Effects of organic farming on field boundary vegetation in Denmark. Agriculture, Ecosystems and Environment, 113, 302–306.

Pfleeger TG, Olszyk D, Burdick C, King G, Kern J and Fletcher J, 2006. Using geographic information systems to identify areas for potential off target pesticide exposures. Environmental Toxicology and Chemistry, 25, 2250–2259.

Pfleeger T, Olszyk D, Lee EH and Plocher M, 2011. Effects of low levels of herbicides on potatoes (*Solanum tuberosum* L.) soybeans (*Glycine max* L.) and peas (*Pisum sataivum* L.) grown in a greenhouse and outside. Environmental Toxicology and Chemistry, 30, 455–468.

Pfleeger T, Blakeley-Smith M, King G, Henry Lee E, Plocher M and Olszyk D, 2012. The effects of glyphosate and aminopyralid on a multi-species plant field trial. Ecotoxicology 21(7), 1–17.

Pocock MJO, Evans DM and Memmott J, 2012. The robustness and restoration of a network of ecological networks. Science, 335, 973–977.

Potts GR, 1980. The effects of modern agriculture, nest predation and game management on the population ecology of partridges (*Perdix perdix* and *Alectoris rufa*). Advances in Ecological Research, 11, 2–79.

Potts GR, Ewald JA and Aebischer NJ, 2010. Long-term changes in the flora of the cereal ecosystem on the Sussex Downs, England, focusing on the years 1968–2005. Journal of Applied Ecology, 47, 215–226.

Prokopy, J R, 2003. Two decades of bottom-up, ecologically based pest management in a small commercial apple orchard in Massachusetts; Agriculture, Ecosystems and Environment 94 (2003) 299–309.

Proulx R, Wirth C, Voigt W, Weigelt A, Roscher C, Attinger S, Baade J, Banard RL, Buchmann N, Buscot F, Eisenhauer N, Fischer M, Gleixner G, Halle S, Hildebrandt A, Kowalski E, Kuu A, Lange M, Milcu A, Niklaus PA, Oelmann Y, Rosenkranz S, Sabais A, Scherver C, Scherer-Lorenzen M, Scheu S, Schulze E-D, Schumacher J, Schwichtenberg G, Soussana J-F, Temperton VM, Weisser WW, Wilcker W and Schmid B, 2010. Diversity promotes temporal stability across levels of ecosystem organization in experimental grasslands. PLoS One, 5, e13382. doi: 10.1371/journal.pone.0013382.

Rautmann D, Streloke M and Winkler R, 2001. New basic drift values in the authorisation procedure for plant protection products. In Workshop on Risk Assessment and Risk Mitigation Measures in the Context of the Authorisation of Plant Protection Products (WORMM). Eds Forster R and Streloke M. Mittelungen aus der Biologischen Bundesanstalt fur Land und Forstwirtschaft, Berlin-Dahlem, Heft 381. 133-141.

Rees M, Grubb PJ and Kelly D, 1996. Quantifying the impact of competition and spatial heterogenity on the structure and dynamics of a four-species guild of winter annuals. The American Naturalist, 147, 1–32.

Reichenberger S, Bach M and Frede HG, 2007. Auswertung der prädiktiven PRZM-Simulationen aus Reichenberger (2005) zum Austrag von PSM-Wirkstoffen über Runoff und Erosion in Deutschland. Dokumentation für den Industrieverband Agrar (IVA). Finale Version vom 10.08.2007. Universitat Gießen.

Reuter S and Siemoneit-Gast S, 2006. Extended method for assessing the risk to terrestrial non-target plants exposed to plant protection products. Poster Presentation. SETAC Europe 16th Annual Meeting, The Hague, 2006.

Reuter S and Simoneit-Gast S, 2007. Entwicklung einer weiterführenden Methode zur Bewertung des Risikos für terrestrische Pflanzen durch Exposition mit Pflanzenschutzmitteln und ihren Wirkstoffen.

Rew LJ, Theaker AJ, Froud-Williams RJ and Boatman ND, 1992. Nitrogen fertilizer misplacement in field boundaries. Aspects of Applied Biology, 30, 203–206.

Rich TCG and Woodruff ER, 1996. Changes in the vascular plant floras of England and Scotland between 1930–1960 and 1987–1988: the BSBI monitoring Scheme. Biological Conservation, 75, 217–229.

Richardson RJ, Wilson HP, Armel GR and Hines TE, 2004. Influence of adjuvants on cotton (*Gossypium hirsutum*) response to postemergence applications of CGA 362622. Weed Technology, 18, 9–15.

Riemens MM, Dueck T and Kempenaar C, 2008. Predicting sublethal effects of herbicides on terrestrial non-crop species in the field from greenhouse data. Environmental Pollution, 155, 141–149.

Riemens MM, Dueck T, Kempenaar C, Lotz LAP and Kropff MJJ, 2009. Sublethal effects of herbicides on the biomass and seed production of terrestrial non-crop plant species, influenced by environment, development stage and assessment date. Environmental Pollution, 157, 2306–2313.

Rinella MJ, Haferkamp MR, Masters RA, Muscha JM, Bellows SE and Vermeire LT, 2010a. Growth regulator herbicides prevent invasive annual grass seed production. Invasive Plant Science and Management, 3,12–16.

Rinella MJ, Masters RA and Bellows SE, 2010b. Growth regulator herbicides prevent invasive annual grass seed production under field conditions. Rangeland Ecology and Management, 63, 487–490.

Roberts HA and Neilson JE, 1981. Changes in the soil seed bank of four long-term crop/herbicide experiments. Journal of Applied Ecology, 18, 661–668.

Rokich DP, Harma J, Turner SR, Sadler RJ and Tan BH, 2009. Fluazifop-p-butyl herbicide: Implications for germination, emergence and growth of Australian plant species. Biological Conservation, 142, 850–869.

Romero A, Chamorro L and Sans FX, 2008. Weed diversity in crop edges and inner fields of organic and conventional dryland winter cereal crops in NE Spain. Agriculture, Ecosystems and Environment, 124, 97–104.

Rose KE, Louda SM, Rees M, 2005). Demographic and evolutionary impacts of native and invasive insect herbivores: a case study with Platte thistle, *Cirsium canescens*. Ecology, 86, 453–465.

Rotchés-Ribalta R, Boutin C, Blanco-Moreno JM, Carpenter DJ and Sans XF, 2012. Effects of herbicide on ecological patterns of characteristic and rare weed species. British Ecological Society, Birmingham, UK.

Rowe L, Rossman E and Penner D, 1990. Differential response of corn hybrids and inbreds to metolachlor. Weed Science, 38, 563–566.

Rowntree K and Sheffield E, 2005. Effects of asulam spraying on non-target ferns. *Canadian Journal of Botany*, 83, 1622–1629.

Rowntree K, Lawton KF, Rumsey FJ and Sheffield E, 2003. Exposure to Asulox inhibits the growth of mosses. Annals of Botany, 92, 547–556.

Rowntree JK, Sheffield E and Burch J, 2005. Growth and development of mosses are inhibited by the common herbicide asulam. The Bryologist, 108, 287–294.SCHER, SCCS, SCENIHR, 2012. Opinion on the toxicity and assessment of chemical mixtures. European Commission 2012.

Schmitz J, Schäfer K and Brühl CA, 2013a. Agrochemicals in field margins—assessing the impacts of herbicides, insecticides, and fertiliser on the common buttercup (*Ranunculus acris*). Environmental Toxicology and Chemistry, 32, 1124–1131.

Schmitz J, Stahlschmidt P and Brühl CA, 2013b. Protection of terrestrial non-target plant species in the regulation of environmental risks to pesticides. Report Gz. Z6 – 93104-4/10.Main Report. FKZ 360 03 053. Available at www.umweltbundesamt.de

Seefeldt SS, Jensen JE and Fürst EP, 1995. Log-logistic analysis of dose–response relationships. Weed Technology, 9, 218–227.

Shea K and Kelly D, 1998. Estimating biocontrol agent impact with matrix models: *Carduus nutans* in New Zealand. Ecological Applications, 8, 824–832.

Siemoneit-Gast S, Reuter S, Kubiak R and Höllrigl-Rosta A, 2007. Development of an extended method for assessing the risk to terrestrial plants exposed to plant protection products and their active ingredients. Editors Re, A. A. M. del;Capri, E.;Fragoulis, G.;Trevisan, M. 13th Symposium Pesticide Chemistry, Piacenza, Italy, 3–6 September 2007. Environmental fate and ecological effects of pesticides 2007 pp. 649–656.

Sikkema PH, Brown L, Shropshire C and Soltani N, 2007. Responses of three types of winter wheat (*Triticum aestivum* L.) to spring-applied post-emergence herbicides. Crop Protection, 26, 715–720.

Silvertown J, Dodd ME, David JG Gowing DJG and Owen Mountford JO, 1999. Hydrologically defined niches reveal a basis for species richness in plant communities. Nature, 400(6739), 61–63.

Smith PG, Boutin C and Knopper L, 2013. Vanadium pentoxide phytotoxicity: effects of species selection and nutrient concentration. Archives of Environmental Contamination and Toxicology, 64, 87–96.

Sotherton NW, Dower JW and Rands NRW, 1988. The effects of pesticide exclusion strips on faunal populations in Great Britain. Ecological Bulletins, 39, 197–199.

Southcombe E, Miller P, Ganzelmeier H, van de Zande J, Miralles A and Hewitt A, 1997. The international (BCPC) spray classification system including drift potential factor. Proceedings Brighton Crop Protection Conference—Weeds, Brighton, UK, 371–280.

Spruijt J, Spoorenberg PM, Rovers JAJM, Slabbekoorn JJ, de Kool SAM and Vlaswinkel MET, 2010. Mogelijkheden om milieubelasting en kosten van gewasbescherming te verlagen, Wettelijke Onderzoekstaken Natuur & Milieu, WOT- report 114.

Stevens CJ, Dise NB, Mountford JO and Gowing DJ, 2004. Impact of nitrogen deposition on the species richness of grasslands. *Science*, **303**, 1876–1879.

Storkey J, Meyer S, Still KS and Leuschner C, 2012. The impact of agricultural intensification and land-use change on the European arable flora. Proceedings of the Royal Society B—Biological Sciences, 279, 1421–1429.

Strandberg B, Mathiassen SK, Damgaard C, Bruus M, 2007. Testing non-target effects of herbicide spray drift. Poster presentation SETAC Europe 17th Annual Meeting, Porto, 20-24 May 2007.

Strandberg B, Mathiassen SK, Bruus M, Kjær C, Damgaard C, Andersen HV, Bossi R, Løfstrøm P, Larsen SE, Bak J and Kudsk P, 2012. Effects of herbicides on non-target plants: How do effects in standard plant tests relate to effects in natural habitats? Pesticide Research No. 137, 2012. Danish Ministry of the Environment, Environmental Protection Agency, 114 pp.

Sullivan TP, Wagner RG, Pitt DG, Lautenschlager RA and Chen DG, 1998. Changes in diversity of plant and small mammal communities after herbicide application in sub-boreal spruce forest. Canadian Journal of Forestry Research, 28, 168–177.

Sutcliffe OL and Kay QON, 2000. Changes in the arable flora of central southern England since 1960s. Biological Conservation, 93, 1–8.

Swarowsky K, Schweikert C, Craig P and Höllrigl-Rosta A, 2014a. How to assess the risk of multiple applications from single application studies? Poster presentation. SETAC Europe 24th Annual Meeting, 11–15 May 2014, Basel, Switzerland.

Swarowsky K, Schweikert K, Craig P, Höllrigl-Rosta A, 2014b. How to derive appropriate limit test application rates for non-target terrestrial plant (NTTP) risk assessment. Poster presentation at the SETAC Europe 24th Annual Meeting, 11-15 May 2014, Basel, Switzerland.

Sweat JK, Horak MJ, Peterson DE, Lloyd RW and Boyer JE, 1998. Herbicide efficacy on four *Amaranthus* species in soybean (*Glycine max*). Weed Technology, 12, 315–321.

Tan WZ, Li QJ and Qing L, 2002. Biological control of alligatorweed (*Alternanthera philoxeroides*) with a *Fusarium* sp. BioControl, 47, 463–479.

Taylor-Lovell S, Wax LM and Nelson R, 2001. Phytotoxic response and yield of soybean (*Glycine max*) varieties treated with sulfentrazone or flumioxazin. Weed Technology, 15, 95–102.

Tew TE, Macdonald DW, Rands MRW, 1992. Herbicide application affects microhabitat use by arable wood mice (*Apodemus sylvaticus*). Journal of Applied Ecology, 29, 532–539.

Thomson WT, 1985. Agricultural chemicals, Book 1: Insecticides, 1985–86 revision. Thomson Publications, Fresno, CA.

Tietjen HP, Halvorson CH, Hegdal PL and Johnson AM, 1967. 2,4-D herbicide, vegetation, and pocket gopher relationships, Black Mesa, Colorado. Ecology, 48, 635–643.

Tiktak A, van den Berg F, Boesten JJTI, van Kraalingen D, Leistra M, van der Linden AMA. 2000. Manual of FOCUS PEARL version 1.1.1. RIVM report 711401008, RIVM, Bilthoven, the Netherlands, pp. 144.

Türe C and Böcük H, 2008. Investigation of threatened arable weeds and their conservation status in Turkey. Weed Research, 48, 289–296.

Umweltbundesamt, 2011. Exposit 3.0 beta version, Available online: www.bvl.bund.de/DE/04_Pflanzenschutzmittel/03_Antragsteller/04_Zulassungsverfahren/07_Naturhaushalt/psm_naturhaush_node.html

US EPA (United States Environmental Protection Agency), 1996. Ecological effects test guidelines: terrestrial plant toxicity—vegetative vigor. OPPTS 850.4150 and 4250. EPA 712-C-96–163, Washington, DC.

US EPA (United States Environmental Protection Agency), 2012a. Ecological effects test guidelines: seedling emergence and seedling growth. OCSPP 850.4100. EPA 712-C-012. Washington, DC.

US EPA (United States Environmental Protection Agency), 2012b. Early seedling growth toxicity test. OCSPP 850.4230. EPA 712-C-010, Washington, DC.

US EPA (United States Environmental Protection Agency), 2012c. Vegetative vigor. OCSPP 850.4150. EPA 712-C-011, Washington, DC.

US EPA (United States Environmental Protection Agency), 2012d. Background and special considerations—tests with terrestrial aquatic plants, cyanobacteria, and terrestrial soil-core microcosms. OCSPP 850.4000. EPA 712-C-013, Washington, DC.

US EPA (United States Environmental Protection Agency), 2012e. Terrestrial plants field study. OCSPP 850.4300. EPA 712-C-009, Washington, DC.

Urwin CP, Wilson RG and Mortensen DA, 1996. Response of dry edible bean (*Phaseolus vulgaris*) cultivars to four herbicides. Weed Technology, 10, 512–518.

Van Calster H, Vandenberghe R, Ruysen M, Verheyen K, Hermy M and Decocq G, 2008. Unexpectedly high 20th century floristic losses in a rural landscape in northern France. Journal of Ecology, 96, 927–936.

Van den Berg F, Doan NK, Jacobs CMJ, Spanoghe P, Butler Ellis MC, Fragkoulis G, 2013. Exposure of workers, residents and bystanders to vapour of plant protection products after application. In: proceedings of the Conference Pesticide Behaviour in Soils, Water and Air, September 2-4, 2013, York, UK, 2pp. Available at http://www.york.ac.uk/conferences/yorkpesticides2013/

Van de Zande JC, Holterman HJ, Huijsmans JFM (2012). Spray drift for the assessment of aquatic organisms to plant protection products in the Netherlands. Part I. Field crops and downward spraying. Wageningen UR, Plant Research International Report 419, Wageningen, the Netherlands.

Van de Zande JC, Rautmann D, Holterman HJ, Huijsmans JFM (2014). Joined spray drift curves for boom sprayers in The Netherlands and Germany. Wageningen UR, Plant Research International Report 526, Wageningen, the Netherlands.

Van Eerdt MM, Spruijt J, van Zeijts and Tiktak A, 2014. Costs and effectiveness of on-farm measures to reduce aquatic risks from pesticides in the Netherlands. Pest Manag Sci. 2014 Jan 20. doi: 10.1002/ps.3729

Van Jaarsveld JA, 2004. The Operational Priority Substances Model. Description and validation of OPS-Pro 4.1. RIVM report 500045001/2004, RIVM Bilthoven, The Netherlands, pp. 156.

Van Vlaardingen PLA, Traas TP, Wintersen AM and Aldenberg T. ETX 2.0 A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. RIVM report 601501028/2004.

Walker ER and Oliver LR, 2008. Weed seed production as influenced by glyphosate applications at flowering across a weed complex. Weed Technology, 22, 318–325.

Walker KJ, Critchley CNR, Sherwood AJ, Large R, Nuttall P, Hulmes S, Rose R and Mountford JO, 2007. The conservation of arable plants on cereal field margins: an assessment of new agri-environment scheme options in England, UK. Biological Conservation, 136, 260–270.

Wang W and Freemark K, 1995. The use of plants for environmental monitoring and assessment. Ecotoxicology and Environmental Safety, 30, 289–301.

Wax LM, Stoller EW and Bernard RL, 2006. Differential response of soybean cultivars to metribuzin. Agronomy Journal, 68, 484–486.

Weidenhamer JD, Hartnett DC and Romeo JT, 1989. Density-dependent phytotoxicity distinguishing resource competition and allelopathic interference in plants. Journal of Applied Ecology, 26, 613–624.

Weiher E, Clarke GDP and Keddy PA, 1998. Community assembly rules, morphological dispersion, and the coexistence of plant species. Oikos, 81, 309–322.

Weiner J, 1990. Asymmetric competition in plant populations. Trends in Evolution and Ecology, 5, 360–364.

Weisser P, Landfried M and Koch H, 2002. Off-crop drift sediments on plant surfaces—exposure of non-target organisms. Aspect of Applied Biology, 66, 225–230.

Westoby M and Wright J, 2006. Land-plant ecology on the basis of functional traits. Trends in Evolution and Ecology, 21, 261–268.

White AL and Boutin C, 2007. Herbicidal effects on non-target vegetation: investigating the limitations of current pesticide registration guidelines. Environmental Toxicology and Chemistry, 26, 2634–2643.

White AL, Boutin C, Dalton RL, Henkelman B and Carpenter D, 2009. Germination requirements for 29 terrestrial and wetland wild plant species appropriate for phytotoxicity testing. Pest Management Science, 65, 19–26.

Wilkinson RE, 1980. Ecotypic variation of *Tamarix pentandra* epicuticular wax and possible relationship with herbicide sensitivity. Weed Science, 28, 110–113.

Wilkinson RE and Mayeux HS, 1987. Composition of epicuticular wax on *Isocoma* leaves. Botanical Gazette, 148, 12–16.

Wilson PJ, 1994. Managing field margins for the conservation of the arable flora. British Crop Protection Council Monograph, 58, 253–258.

Wilson PJ and Aebischer NJ, 1995. The distribution of dicotyledonous arable weeds in relation to distance from the crop edge. Journal of Applied Ecology, 32, 295–310.

Wilson PJ and King M, 2003. Arable plants—a field guide. English Nature and Wildlife Guides, Old Basing, Hampshire, UK.

Winkle ME, Leavitt JRC and Burnside OC, 1981. Effects of weed density on herbicide absorption and bioactivity. Weed Science, 29, 405–409.

Wyrill JB and Burnside OC, 1976. Absorption, translocation, and metabolism of 2,4-D and glyphosate in common milkweed and hemp dogbane. Weed Science, 24, 557–566.

Zhou J, Tao B, Messersmith CG and Nalewaja JD, 2007. Glyphosate efficacy on velvetleaf (*Abutilon theophrasti*) is affected by stress. Weed Science, 55, 240–244.

Zwerger P and Pestemer W, 2000. Testing the phytotoxic effects of herbicides on higher terrestrial non-target plants using a plant life cycle test. Zeitschrift für Pflanzenkranheiten und Pflanzenschutz Sonderh, 17, 711–718.

APPENDICES

Appendix A. Extrapolation from ER_{50}s for vegetative endpoints to ER_{10}s for reproductive endpoints

A.1. Introduction

In Table 2 about specific protection goals in section 2 of this opinion a number of endpoints were proposed for use in risk assessment for non-target terrestrial plants. For some protection goals the endpoint is a 5th percentile of the ER_{10} (effective concentration resulting in a 10 % decrease compared with the controls) for reproduction, used as a surrogate for no observable effect rates (NOERs), for others it is the 5th percentile of the ER_{10} for biomass, the 5th percentile of the ER_{50} values for biomass or the 5th percentile of the ER_{50} values for visual endpoints. The data available do not always enable all three endpoints to be derived; in such cases, extrapolation between the tested endpoints and the required ones are necessary. When the original data from the studies are suitable to calculate these values, those will be the preferred ones to use in the risk assessment for non-target plants. According to Commission Regulation (EU) No 283/2013, for active substances that exhibit herbicidal or plant growth regulator activity, vegetative vigour and seedling emergence concentration/response tests shall be provided. It is further stated that dose–response tests on a selection of 6 to 10 monocotyledon and dicotyledon plant species representing as many taxonomic groups as possible shall be provided. It is also stated in the regulation that the ER_{10}, ER_{20} and ER_{50} shall be reported together with the NOER (section 8, Introduction, point 6).

The standard endpoint from the plant toxicity tests is an ER_{50} value for vegetative endpoints. Because it is intended to protect plant populations it is advisable to use reproductive endpoints and to use an endpoint at which no effects were observed (NOER). An ER_{10} value will be used as a surrogate NOER value (see notes under Table 2). Note that the reproductive endpoint is not always the lowest available endpoint and that the plant species diversity can also be influenced by indirect effects, e.g. due to competitive interactions in the field, too, which are not covered by the reproductive endpoint.

A.2. Methods

For nine herbicides (2,4-D, chlorimuron ethyl, glufosinate ammonium, glyphosate, mecoprop, metsulfuron methyl, primisulfuron, sulfometuron and tribenuron) first the standard ER_{50} for a vegetative endpoint was assessed and thereafter the study was prolonged to obtain a reproductive endpoint. In total 55 tests (34 different species) were available for which a vegetative endpoint and a reproductive endpoint were also available (see Table A1).

A.2.1. Test design

In Carpenter and Boutin (2010), plants (six replicates) were exposed to nine doses of glufosinate ammonium ranging from 0 to 667.5 g a.i./ha^{-1} (0–89 % of label rate) at the three- to six-leaf stage. In Carpenter et al. (2013), six plant replicates were exposed to eight doses of chlorimuron ethyl ranging from 0 to 9.63 g a.i./ha (0–107 % of label rate) at the four- to six-leaf stage. In Strandberg et al. (2012) and Mathiassen (unpublished data), seeds were directly sown in 2-L pots or were sown in trays and transplanted into pots as small seedlings. Herbicides were applied at the four- to eight-leaf stage. Pots were placed outdoors following the herbicide exposure and stayed there until plants for biomass measurements were selected, i.e. three to four weeks after exposure. Thereafter, the pots were again moved to the greenhouse (controlled conditions) to optimise conditions for seed production during the autumn period. Honey bees were available in the greenhouse to ensure pollination. Pots were watered several times daily. The experimental work in Rotchés-Ribalta et al. (2012) included seven doses and five replicates with herbicide doses ranging from 0 to 7.5 g a.i./ha for tribenuron and 0 to 564 g a.i./ha for 2,4-D, the equivalent of 0 % to 100 % of label rate in both cases. In Olszyk et al. (2009), six plant replicates were sprayed with five doses ranging from 0 % to 10 % label rate, corresponding to 0.04 to 4 g a.i./ha for primisulfuron, 0.053 to 5.3 g a.i./ha for sulfometuron and 0.833 to 83.3 g a.i./ha for

glyphosate. Except for Strandberg et al. (2012) and Olszyk et al. (2009), experiments were conducted under greenhouse conditions. Herbicide and species used, as well as growth stage at time of spraying, test duration and endpoints measured are presented in Table A1.

A.2.2. ER_{10} and ER_{50} calculations

In experiments conducted by Carpenter Boutin (2010) Carpenter et al. (2013) and Rotchés-Ribalta et al. (2012), the ER_{10} and ER_{50} were calculated using non-linear regressions when the data met the assumptions of normality and homogeneity of variance, or else the non-parametric ICPIN method was used (Norberg-King, 1993). Vegetative and reproductive parameters (seed production or measurable equivalent) were used separately in each calculation. In Olsyk et al. (2009), ER_{10} and ER_{50} were recalculated with the raw data provided by the authors. Similarly in Strandberg et al. (2012) and Mathiassen (unpublished), the ERs were analysed with non-linear regressions using log-logistic dose–response models (Seefeldt et al., 1995). For each herbicide, dose–response curves were estimated for each plant species and growth stage. The fitness of the model was verified using an *F*-test for lack of fit, comparing the residual sum of squares.

A.3. Results

In some cases the reproductive endpoint is greater than the vegetative endpoint; when comparing ER_{50} values for a vegetative endpoint with an ER_{50} for a reproductive endpoint, 16 out of the possible 39 combinations (41 %) show a lower vegetative endpoint than that for reproduction. When comparing ER_{10} values for a vegetative endpoint with an ER_{10} for a reproductive endpoint, 20 out of the possible 50 combinations (40 %) provide a lower vegetative endpoint than that for reproduction. When comparing an ER_{50} for a vegetative endpoint with an ER_{10} for a reproductive endpoint, the latter is always lower except for 2 out of 42 combinations (2 %). The average ratio between the ER_{50} vegetative endpoint and the ER_{10} reproductive endpoint is 36 (42 combinations). The 50^{th} percentile of the outcomes is 5.3, and the 75^{th}, 90^{th} and 95^{th} percentiles are 25, 56 and 184, respectively. All outcomes of the ER_{50} vegetative endpoint divided by the ER_{10} reproductive endpoint are presented in Figure A1.

A.3.1. Extrapolation from an ER_{10} vegetative endpoint (ER_{veg10}) to an ER_{10} reproductive endpoint ($ER_{repro10}$)

To use this extrapolation method, the dose–response curves for the individual tested plant species need to be available. For each plant species the ER_{10} is assessed. For each compound with six or more values for individual species, the species sensitivity distribution (SSD) approach to calculate the 5^{th} percentile of the distribution of effect rates (e.g. the HR_5 of ER_{veg10}) is used and the endpoint is then associated with an EF.

For the compounds with fewer than six values for individual species, the lowest available value is divided by 5 to derive a surrogate HC_5. The results are presented in Table A2.

When an EF of 2 is used, 3 out 48 comparisons will not be covered by the EF, which is 6.3 %. For EFs of 3 and 5, the percentages are 2.1 % and 0 %, respectively. To achieve a covering of 95 % of the comparisons an EF of 3 has to be used. When all comparisons should be covered, an EF of 5 has to be implemented.

A.3.2. Extrapolation from an ER_{50} vegetative endpoint (ER_{veg50}) to an ER_{10} reproductive endpoint ($ER_{repro10}$)

This extrapolation method can be used when no information is available for calculating the ER_{veg10} from the dose–response curve. The same approach as described in the section above is used and the calculated values are compared with the available $ER_{repro10}$ values to calculate an appropriate EF factor. The results of this assessment are presented in Table A3.

When an EF of 5 is used, 11 out 48 comparisons will not be covered by the EF, which is 23 %. For EFs of 10, 20 and 30 the percentages are 19 %, 10 % and 8 %, respectively. For an EF of 40 this percentage is 2 %. To achieve a covering of 95 % of the comparisons, an EF of 35 has to be used. When all comparisons should be covered, an EF of 70 has to be implemented.

Alternatively, for a part of the active substances of the dataset for which appropriate data are available, the fraction affected from the respective $ER_{repro10}$ SSD can be calculated for each substance using the actual HR_5/EF as a basis (see Table A5). As in Tables A2 to A4 the lowest endpoint/5 was used as a surrogate HR_5 if fewer than six species were available to calculate the HR_5 of the ER_{50} for the vegetative endpoint.

Using the SSD approach, for EFs of 5, 10, 20, 30, 40, 50, 70 and 85 the fractions affected for a set of individual substances were calculated. The highest values correspond to the substance 2,4-D and are 49, 33, 20, 14, 11, 9, 6, and 5 %, respectively. On the basis of Table A5, using an EF of 35 will be protective for six out of nine substances evaluated.

A.3.3. Extrapolation from an ER_{50} vegetative endpoint (ERv_{eg50}) to an ER_{10} vegetative endpoint (ER_{veg10})

This extrapolation method can be used when no information is available for calculating the ER_{veg10} from the dose–response curve and the risk assessment is based on the 5th percentile of the ER_{veg10} values. The same approach has been used as above and the outcome is presented in Table A4.

When an EF of 10 is used, 10 out of 48 comparisons will not be covered by the EF, which is 21 %. For EFs of 20, 30 and 40, the percentages are 10 %, 6 % and 2 %, respectively. To achieve a covering of 95 % of the comparisons, an EF of 34 has to be used. When all comparisons should be covered an EF of 350 has to be implemented.

A.3.4. Remarks

It is important to note that the endpoint selected (e.g. HR_5 of SSD) will then have to be associated during the risk assessment with an AF that will cover for the remaining uncertainties (e.g. single species to multispecies (ecological interactions), environmental stressors).

The EFs defined here are associated with some inherent uncertainties owing to the nature of the data used in this exercise. Only a few studies were available and some are carried out by the same authors. It is advisable to redo these calculations when more data are available in future and to pay more attention to the representativeness of the test species and the potential grouping of species (e.g. annual versus perennial species).

An HR_5 based on ER_{10}s will introduce an additional source of uncertainty. The uncertainty around an ER_{50} is smaller than the uncertainty around an ER_{10}. Please note that, on the one hand, ER_{10} is a relevant endpoint as it takes into consideration the slope of the dose–response curve. But, on the other hand, defining an ER_{10} is linked to more uncertainties than an ER_{50} (statistically less robust).

The appropriateness of using a factor X (in this case 5) for small datasets needs to be verified.

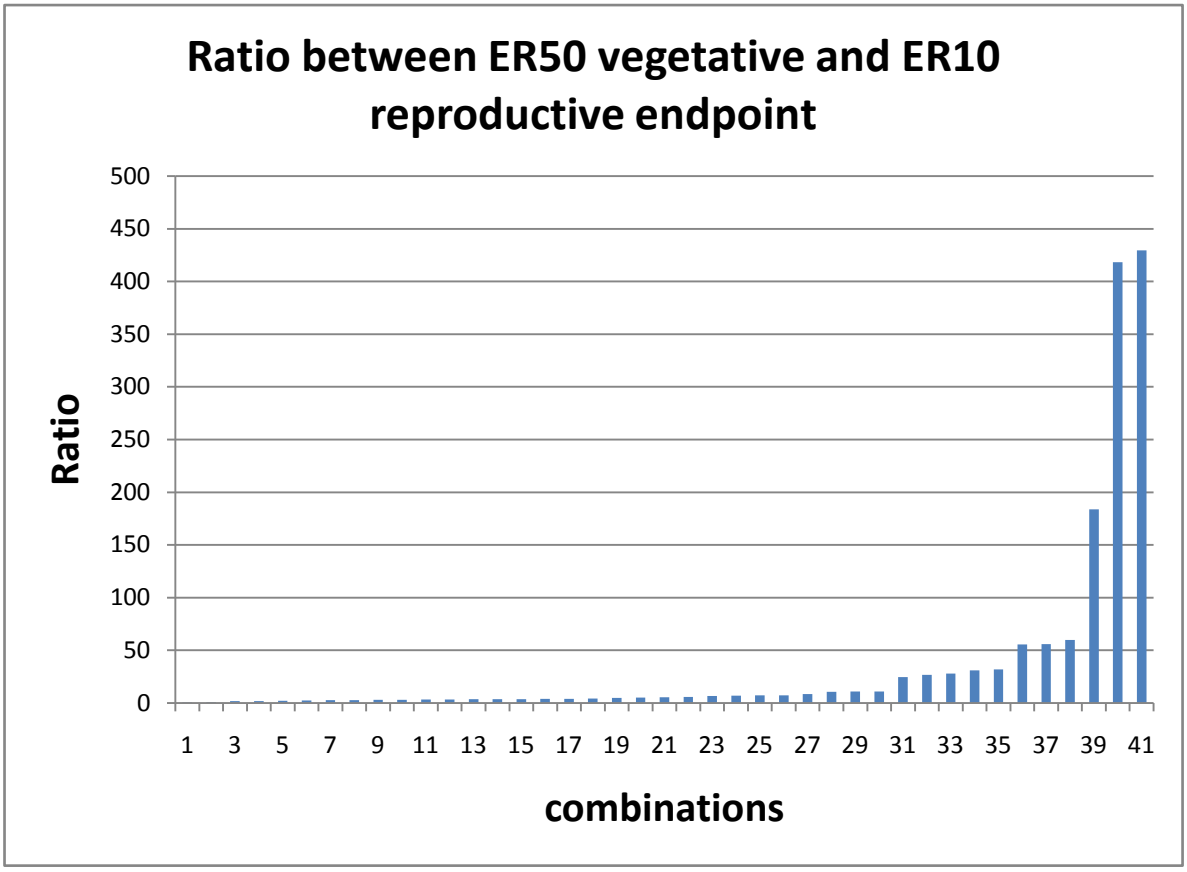


Figure A1: Ratio between vegetative endpoint (ER_{50}) and reproductive endpoint (ER_{10}) for 42 individual studies (see Table A1**)**

Table A1: Toxicity tests for which ER_{50}s and/or ER_{10}s are available for vegetative and reproductive endpoints

Compound	Species	Growth stage or days after emergence	Test duration (time after spray)		Endpoint measured		Vegetative juvenile		Reproduction		Reference
			Vegetative	Reproductive	Vegetative	Reproductive	ER_{50}	ER_{10}	ER_{50}	ER_{10}	
Chlorimuron ethyl	*Capsella bursa-pastoris*	4–6 leaf stage	28 days	50 days	Above ground dry biomass	Number of pods	1.53	0.33	0.66	0.22	Carpenter et al., 2013
	Centaurea cyanus	4–6 leaf stage	28 days	99 days	Above ground dry biomass	Number of flower heads	6.95	0.44	-	0.65	Carpenter et al., 2013
	Chenopodium album	4–6 leaf stage	28 days	57 days	Above ground dry biomass	Number of seeds	-	0.84	6.65	0.18	Carpenter et al., 2013
	Helianthus strumosus	4–6 leaf stage	28 days	116 days	Above ground dry biomass	Number of seedhead florets	1.85	0.61	2.49	0.57	Carpenter et al., 2013
	Lobelia inflata	4–6 leaf stage	28 days	92 days	Above ground dry biomass	Number of fruits	0.66	0.10	3.74	0.86	Carpenter et al., 2013
	Anagallis arvensis	4–6 leaf stage	28 days	131 days	Above ground dry biomass	Number of fruits	-	0.54	1.92	0.03	Carpenter et al., 2013
	Glyceria striata	4–6 leaf stage	28 days	85 days	Above ground dry biomass	Tiller count	0.63	0.11	1.54	0.89	Carpenter et al., 2013
	Lycopus americana	4–6 leaf stage	28 days	99 days	Above ground dry biomass	Floral nodes	2.61	0.29	3.59	0.71	Carpenter et al., 2013
	Polygonum pensylvanicum	4–6 leaf stage	28 days	119 days	Above ground dry biomass	Number of seeds	1.67	0.20	3.36	0.48	Carpenter et al., 2013
Glufosinate ammonium	*Avena sativa*	3–6 leaf stage	21 days	59 days	Above ground dry biomass	Seed production	216.77	39.55	149.31	31.89	Carpenter and Boutin., 2010
	Fagopyrum esculentum	3–6 leaf stage	21 days	106 days	Above ground dry biomass	Seed production	56.02	11.56	113.6	2.01	Carpenter and Boutin., 2010
	Helianthus annuus	3–6 leaf stage	21 days	101 days	Above ground dry biomass	Seedhead mass	117.3	72.96	145.25	64.30	Carpenter and Boutin., 2010
	Solanum lycopersicum	3–6 leaf stage	21 days	127 days	Above ground dry biomass	Fresh fruit weight	65.37	17.16	145.89	15.07	Carpenter and Boutin., 2010
	Bouteloua gracilis	3–6 leaf stage	21 days	78 days	Above ground dry	Tiller count	115.95	15.67	101.09	10.67	Carpenter and Boutin., 2010

Compound	Species	Growth stage or days after	Test duration (time after spray)		Endpoint measured		Vegetative juvenile		Reproduction		Reference
					biomass						
	Elymus canadensis	3–6 leaf stage	21 days	100 days	Above ground dry biomass	Seed production	165.04	4.29	43.08	28.16	Carpenter and Boutin., 2010
	Juncus dudleyi	3–6 leaf stage	21 days	72 days	Above ground dry biomass	Number of fruits	154.31	53.58	49.11	28.92	Carpenter and Boutin., 2010
	Capsella bursa-pastoris	3–6 leaf stage	21 days	38 days	Above ground dry biomass	Number of pods	33.37	8.05	41.49	10.48	Carpenter and Boutin., 2010
	Hypericum perforatum	3–6 leaf stage	21 days	139 days	Above ground dry biomass	Apical meristem	81.68	16.82	40.99	27.88	Carpenter and Boutin., 2010
	Melilotus officinalis	3- to 6-leaf stage	21 days	66 days	Above ground dry biomass	Seed production	36.08	5.19	31.49	1.13	Carpenter and Boutin., 2010
	Phytolacca americana	3- to 6-leaf stage	21 days	101 days	Above ground dry biomass	Number of fruits	97.17	53.33	62.74	1.75	Carpenter and Boutin., 2010
	Solanum dulcamara	3- to 6-leaf stage	21 days	125 days	Above ground dry biomass	Meristem	40.68	25.93	94.28	15.01	Carpenter and Boutin., 2010
Mecoprop	*Silene noctiflora*	6–8 leaves	3–4 weeks	At maturity	Fresh weight	Number of seeds	69	7.3	38.1	8.1	Strandberg et al., 2012
	Silene vulgaris	6–8 leaves	3–4 weeks	At maturity	Fresh weight	Number of seeds	154	13.5	-	-	Strandberg et al., 2012
	Geranium molle	6-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	137.1	41.9	-	37.5	Strandberg et al., 2012
	Geranium robertianum	6-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	54.6	0.7	-	-	Strandberg et al., 2012
Glyphosate	*Silene noctiflora*	6- to 8-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	74.4	25.4	87.2	39.2	Strandberg et al., 2012
	Silene vulgaris	6- to 8-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	70.8	21.3	37.6	17.5	Strandberg et al., 2012
	Geranium molle	6-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	22.9	5.1	-	-	Strandberg et al., 2012
	Geranium robertianum	6-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	108.2	19.6	-	45	Strandberg et al., 2012
	Echinochloa crus-galli	4-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	86.3	60.7	175	22.5	Mathiassen unpublished
	Echinochloa crus-galli	6- to 8-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	44.4	16.4	46.6	20	Mathiassen unpublished
Metsulfuron methyl	*Silene noctiflora*	6- to 8-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	0.6	0.1	0.34	0.12	Strandberg et al., 2012

Compound	Species	Growth stage or days after	Test duration (time after spray)		Endpoint measured		Vegetative juvenile		Reproduction		Reference
	Silene vulgaris	6- to 8-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	-	-	1	0.42	Strandberg et al., 2012
	Geranium molle	6-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	0.7	0.01	-	-	Strandberg et al., 2012
	Geranium robertianum	6-leaf stage	3–4 weeks	At maturity	Fresh weight	Number of seeds	0.33	0.05	0.25	0.125	Strandberg et al., 2012
Tribenuron	*Rapistrum rugosum*	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	0.51	0.08	0.84	0.10	Rotchés-Ribalta et al., 2012
	Galium aparine (spurium)	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	-	5.93	-	0.32	Rotchés-Ribalta et al., 2012
	Papaver rhoeas	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	0.93	0.10	0.17	0.03	Rotchés-Ribalta et al., 2012
	Papaver argemone	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	0.18	0.03	0.25	0.05	Rotchés-Ribalta et al., 2012
	Scandix pecten-veneris	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	4.636	0.256	1.5	0.641	Rotchés-Ribalta et al., 2012
	Buplureum rotundifolium	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	1.188	0.189	0.383	0.048	Rotchés-Ribalta et al., 2012
2,4-D	*Rapistrum rugosum*	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	189.11	44.81	84.17	3.38	Rotchés-Ribalta et al., 2012
	Neslia paniculata	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	204.12	50.29	109.51	18.57	Rotchés-Ribalta et al., 2012
	Galium aparine (spurium)	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	-	1.00	-	92.24	Rotchés-Ribalta et al., 2012
	Spergula arvensis	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	-	516.49	-	14.60	Rotchés-Ribalta et al., 2012
	Papaver rhoeas	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	-	18.99	402.00	1.38	Rotchés-Ribalta et al., 2012
	Papaver argemone	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	480.95	53.45	69.15	1.15	Rotchés-Ribalta et al., 2012
	Scandix pecten-veneris	4- to 6-leaf stage	1 month	2 months	Above ground dry biomass	Seed production	206.49	86.10	79.04	27.79	Rotchés-Ribalta et al., 2012

Compound	Species	Growth stage or days after	Test duration (time after spray)		Endpoint measured		Vegetative juvenile		Reproduction		Reference
Primisulfuron	*Pisum sativum*	14 days after emergence, spring	14 days	35 days	Height	Pea dry weight	3.406	0.012	0.240	0.057	Olszyk et al., 2009
	Pisum sativum	14 days after emergence, summer	14 days	35 days	Height	Pea dry weight	2.899	0.161	0.233	0.108	Olszyk et al., 2009
Sulfometuron	*Pisum sativum*	14 days after emergence, spring	14 days	35 days	Height	Pea dry weight	2.148	0.074	0.038	0.005	Olszyk et al., 2009
	Pisum sativum	14 days after emergence, summer	14 days	35 days	Height	Pea dry weight	1.655	0.023	0.064	0.009	Olszyk et al., 2009
Glyphosate	*Pisum sativum*	14 days after emergence, spring	14 days	35 days	Height	Pea dry weight	-	0.346	27.087	10.612	Olszyk et al., 2009
	Pisum sativum	14 days after emergence, summer	14 days	35 days	Height	Pea dry weight	-	27.708	24.757	8.972	Olszyk et al., 2009

Table A2: Calibration of extrapolation factor (EF) for achieving the protection goal for 95 % of higher plant species. Extrapolation from an ER_{10} vegetative endpoint (ER_{veg10}) to an ER_{10} reproductive endpoint ($ER_{repro10}$)

A	Chlorimuron ethyl	Glufosinate ammonium	Mecoprop	Glyphosate	Tribenuron	2,4-D	Primisulfuron	Sulfometuron	Metsulfuron methyl	Total	%
HR_5 of ER_{10} for vegetative endpoint	**0.087** [(a)]	**3.9** [(a)]	**0.14** [(b)]	**0.64** [(a)]	**0.0084** [(b)]	**1.42** [(a)]	**0.0024** [(a)]	**0.0046** [(b)]	**0.002** [(b)]		
EF = 2	0.043	1.95	0.07	0.32	0.0042	0.71	0.0012	0.0023	0.001		
EF = 3	0.029	1.30	0.046	0.21	0.0028	0.47	0.0008	0.0015	0.0007		
EF = 5	0.017	0.78	0.028	0.13	0.0017	0.28	0.00048	0.0009	0.0004		
B	**Number of ER_{10}s for reproductive endpoint below 5th percentile (or lowest/5)/EF**										
EF = 2	1	2	0	0	0	0	0	0	0	3	**6.3**
EF = 3	0	1	0	0	0	0	0	0		1	**2.1**
EF = 5	0	0	0	0	0	0	0	0	0	0	**0.0**
Number of ER_{10}s for reproductive endpoint											
	9	12	2	5	6	7	2	2	3	48	

(a): Fifth percentile approach.
(b): Lowest available toxicity value divided by 5.

Table A3: Calibration of extrapolation factor (EF) for achieving the protection goal for 95 % of higher plant species. Extrapolation from an ER_{50} vegetative endpoint (ER_{veg50}) to an ER_{10} reproductive endpoint ($ER_{repro10}$)

A	Chlorimuron ethyl	Glufosinate ammonium	Mecoprop	Glyphosate	Tribenuron	2,4-D	Primisulfuron	Sulfometuron	Metsulfuron methyl	Total	%
HR_5 of ER_{50} for vegetative endpoint	**0.4** [(a)]	**28.9** [(a)]	**10.9** [(b)]	**22.7** [(a)]	**0.036** [(b)]	**40.8** [(b)]	**0.58** [(b)]	**0.33** [(b)]	**0.066** [(b)]		
EF = 5	0.080	5.78	2.18	4.54	0.0072	8.16	0.12	0.066	0.0132		
EF = 10	0.040	2.89	1.09	2.27	0.0036	4.08	0.06	0.033	0.0066		
EF = 20	0.020	1.45	0.55	1.14	0.0018	2.04	0.03	0.017	0.0033		
EF = 30	0.013	0.96	0.36	0.76	0.0012	1.36	0.02	0.011	0.0022		
EF = 40	0.010	0.72	0.27	0.57	0.0009	1.02	0.01	0.008	0.0017		
EF = 50	0.008	0.58	0.22	0.45	0.0007	0.82	0.01	0.007	0.0013		
EF = 70	0.006	0.413	0.156	0.324	0.001	0.583	0.008	0.0047	0.0009		
B	**Number of ER_{10}s for reproductive endpoint below 5th percentile (or lowest/5)/EF**										
EF = 5	1	3	0	0	0	3	2	2	0	11	**23**
EF = 10	1	3	0	0	0	3	0	2	0	9	**19**
EF = 20	0	1	0	0	0	2	0	2	0	5	**10**
EF = 30	0	0	0	0	0	2	0	2	0	4	**8**
EF = 40	0	0	0	0	0	0	0	1	0	1	**2**
EF = 50	0	0	0	0	0	0	0	1	0	1	**2**
EF = 70	0	0	0	0	0	0	0	0	0	0	**0**
Number of ER_{10}s for reproductive endpoint											
	9	12	2	5	6	7	2	2	3	48	

(a): Fifth percentile approach.
(b): Lowest available toxicity value divided by 5.

Table A4: Calibration of extrapolation factor (EF) for achieving the protection goal for 95 % of the higher plant species. Extrapolation from an ER_{50} vegetative endpoint (ER_{veg50}) to an ER_{10} vegetative endpoint (ER_{veg10})

A	Chlorimuron ethyl	Glufosinate ammonium	Mecoprop	Glyphosate	Tribenuron	2,4-D	Primisulfuron	Sulfometuron	Metsulfuron methyl	Total	%
HR_5 of ER_{50} for vegetative endpoint	**0.4** [(a)]	**28.9** [(a)]	**10.9** [(b)]	**22.7** [(a)]	**0.036** [(b)]	**40.8** [(b)]	**0.58** [(b)]	**0.33** [(b)]	**0.066** [(b)]		
EF = 10	0.04	2.89	1.09	2.27	0.0036	4.08	0.058	0.033	0.0066		
EF = 20	0.02	1.45	0.55	1.135	0.0018	2.04	0.029	0.0165	0.0033		
EF = 30	0.013	0.96	0.36	0.757	0.0012	1.36	0.019	0.011	0.0022		
EF = 40	0.01	0.72	0.27	0.567	0.0009	1.02	0.0145	0.0083	0.00165		
EF = 350 (*circa*)	0.0011	0.083	0.03	0.0649	0.0001	0.12	0.0016	0.00094	0.00019		
B	**Number of ER_{10}s for vegetative endpoint below 5th percentile (or lowest/5)/EF**										
EF = 10	1	3	0	0	0	3	1	2	0	10	**20.8**
EF = 20	0	1	0	0	0	2	0	2		5	**10.4**
EF = 30	0	0	0	0	0	1	0	2	0	3	**6.3**
EF = 40	0	0	0	0	0	0	0	1	0	1	**2.1**
EF = 350	0	0	0	0	0	0	0	0	0	0	**0**
Number of ER_{10}s for reproductive endpoint											
	9	12	2	5	6	7	2	2	3	48	

(a): Fifth percentile approach.
(b): Lowest available toxicity value divided by 5.

Table A5: Calibration of extrapolation factor (EF) for achieving the protection goal for 95 % of higher plant species. Extrapolation from an ER_{50} vegetative endpoint (ER_{veg50}) to an ER_{10} reproductive endpoint ($ER_{repro10}$)

A	Chlorimuron ethyl	Glufosinate ammonium	Mecoprop	Glyphosate	Tribenuron	2,4-D	Primisulfuron	Sulfometuron	Metsulfuron methyl
HR_5 of ER_{50} for vegetative endpoint	**0.4 [a]**	**28.9 [a]**	**10.9 [b]**	**22.7 [a]**	**0.036 [b]**	**40.8 [b]**	**0.58 [b]**	**0.33 [b]**	**0.066 [b]**
EF = 5	0.080	5.78	2.18	4.54	0.0072	8.16	0.12	0.066	0.0132
EF = 10	0.040	2.89	1.09	2.27	0.0036	4.08	0.06	0.033	0.0066
EF = 20	0.020	1.45	0.55	1.14	0.0018	2.04	0.03	0.017	0.0033
EF = 30	0.013	0.96	0.36	0.76	0.0012	1.36	0.02	0.011	0.0022
EF = 40	0.010	0.72	0.27	0.57	0.0009	1.02	0.015	0.008	0.0017
EF = 50	0.008	0.58	0.22	0.45	0.0007	0.82	0.012	0.007	0.0013
EF = 70	0.006	0.413	0.156	0.324	0.0005	0.583	0.008	0.0047	0.0009
EF = 85	0.005	0.34	0.128	0.267	0.0004	0.48	0.007	0.004	0.0008
B	**Fraction affected based on reproductive endpoint median (LCL – UCL)**								
			No SSD calculable				No SSD calculable	No SSD calculable	No SSD calculable
EF = 5	9.1 (1.6–28.3)	30.5 (15.9–49)	×	< 0.01 (< 0.01–4.95)	1.78 (0.03–19.25)	48.55 (25.51–72.09)	×	EF = 330 required to cover surrogate HC_5 for reproductive endpoint (seed production), i.e. lowest available EC_{10}/5	√
EF = 10	2.6 (0.2–15.8)	14.64 (5.05–31.54)	×	< 0.1	0.42 (< 0.01–11.98)	32.96 (13.56–58.49)	×		√
EF = 20	0.5 (< 0.1–8)	5.59 (1.02–18.4)	√	< 0.1	< 0.1	19.96 (5.58–45.64)	×		√
EF = 30	0.2 (< 0.1–4.9)	2.8 (0.31–12.67)	√	< 0.1	< 0.1	14.07 (2.58–38.85)	×		√
EF = 40	0.1 (< 0.1–3.56)	1.63 (0.12–9.56)	√	< 0.1	< 0.1	10.7 (1.1–34.41)	×		√
EF = 50	< 0.1	1.3 (0.07–9.05)	√	< 0.1	< 0.1	8.56 (0.35–31.26)	×		√

A	Chlorimuron ethyl	Glufosinate ammonium	Mecoprop	Glyphosate	Tribenuron	2,4-D	Primisulfuron	Sulfometuron	Metsulfuron methyl
EF = 70	< 0.1	0.66 (< 0.1–6.44)	√	< 0.1	< 0.1	5.88 (< 0.1–26.73)	√		√
EF = 85	< 0.1	0.43 (< 0.1–5.25)	√	< 0.1	< 0.1	4.68 (< 0.1–24.35)	√		√

(a): Fifth percentile approach.
(b): Lowest available toxicity value divided by 5.
√, surrogate HR_5 for vegetative endpoint/EF covers surrogate HR_5 for reproductive endpoint (seed production), i.e. lowest available $ER_{10}/5$.
LCL, lower confidence level; UCL, upper confidence level.

Appendix B. Comparison of LOER for vegetative parts and LOER for reproduction determined after exposure during the reproductive stage

Reproduction measured as seed production

Table B1: Summary of effects when plants were sprayed during reproductive stages and reproduction assessed. The factor represents the ratio of LOER for vegetative parts to LOER for reproduction.

Reference	Species and family	Herbicide	Phenological stage at spraying	LOER for vegetative parts	LOER for reproduction	Factor
Fletcher et al., 1993	*Prunus avium*, var. Royal Anne	Chlorsulfuron	Flower stage	2.3×10^{-6} M	2.3×10^{-5} M	0.10
			Post-flower stage	2.3×10^{-6} M	4.7×10^{-7} M	4.89
			Small fruit stage	2.3×10^{-6} M	2.3×10^{-6} M	1.00
			Full size stage	2.3×10^{-6} M	2.3×10^{-7} M	10.00
Fletcher et al., 1995	*Pisum sativum*	Chlorsulfuron	Bud stage	4.6×10^{-2} g a.i./ha	9.2×10^{-2} g a.i./ha	0.50
			Open flower stage	1.8×10^{-1}g a.i./ha	1.8×10^{-1} g a.i./ha	1.00
Fletcher et al., 1996	*Brassica napus*	Chlorsulfuron	Pre-flower	9.2×10^{-5} kg a.i./ha	4.6×10^{-5} kg a.i./ha	2.00
			Flowering	4.6×10^{-5} kg a.i./ha	4.6×10^{-5} kg a.i./ha	1.00
			Late flowering	$> 9.2 \times 10^{-5}$ kg a.i./ha	4.6×10^{-5} kg a.i./ha	> 2.00
	Polygonum persicaria	Chlorsulfuron	Pre-flower	4.6×10^{-5} kg a.i./ha	4.6×10^{-5} kg a.i./ha	1.00
			Flowering	1.8×10^{-4} kg a.i./ha	1.8×10^{-4} kg a.i./ha	1.00
			Late flowering	$> 1.8 \times 10^{-4}$ kg a.i./ha	1.8×10^{-4} kg a.i./ha	> 100
	Glycine max	Chlorsulfuron	Pre-flower	9.2×10^{-5} kg a.i./ha	9.2×10^{-5} kg a.i./ha	1.00
			Flowering	$> 1.8 \times 10^{-4}$ kg a.i./ha	4.6×10^{-5} kg a.i./ha	> 391
			Late flowering	$> 1.8 \times 10^{-4}$ kg a.i./ha	9.2×10^{-5} kg a.i./ha	> 196
	Helianthus annuus	Chlorsulfuron	Pre-flower	9.2×10^{-5} kg a.i./ha	1.8×10^{-4} kg a.i./ha	0.51
			Flowering	1.8×10^{-4} kg a.i./ha	$> 1.8 \times 10^{-4}$ kg a.i./ha	1.00
			Late flowering	$> 1.8 \times 10^{-4}$ kg a.i./ha	$> 1.8 \times 10^{-4}$ kg a.i./ha	1.00
	Helianthus annuus	2,4-D	Pre-flower	8.8×10^{-3} kg/ha	8.8×10^{-3} kg/ha	1.00

Reference	Species and family	Herbicide	Phenological stage at spraying	LOER for vegetative parts	LOER for reproduction	Factor
Bhatti et al., 1995	*Prunus avium*, var. Bing	Chlorsulfuron	Full bloom stage	9.3×10^{-7} M	27.9×10^{-7} M	0.33
			Post-bloom stage	9.3×10^{-7} M	27.9×10^{-7} M	0.33
	Prunus avium, var. Chinook	Chlorsulfuron	Full bloom stage	3.1×10^{-7} M	3.1×10^{-7} M	1.00
			Post-bloom stage	3.1×10^{-7} M	3.1×10^{-7} M	1.00
	Prunus avium, var. Rainier	Chlorsulfuron	Full bloom stage	3.1×10^{-7} M	27.9×10^{-7} M	0.11
			Post-bloom stage	9.3×10^{-7} M	27.9×10^{-7} M	0.33
Al-Khatib and Tamhane, 1999	*Pisum sativum*	Chlorsulfuron	Flower buds	0.04 g a.i./ha (0.1 % label rate)	0.18 g a.i./ha (0.7 % label rate)	0.22
		Thifensulfuron	Flower buds	0.09 g a.i./ha (0.3 % label rate)	1.36 g a.i./ha (5.2 % label rate)	0.07
		Dicamba	Flower buds	1.56 g a.i./ha (1 % label rate)	25 g a.i./ha (17.5 % label rate)	0.06
Kjær et al., 2006	*Crataegus monogyna*	Metsulfuron methyl	Bud stage	> 1.6 g a.i./ha (40 % label rate)	0.2 g a.i./ha (5 % label rate)	8.00
			Flower stage	> 1.6 g a.i./ha (40 % label rate)	0.6 g a.i./ha (15 % label rate)	2.67
Boutin et al., 2000	*Mimulus ringens*	Metsulfuron methyl	Flower bud	0.45 g a.i./ha (1 % label rate)	0.45 g a.i./ha (10 % label rate)	1.00
			Onset of flowering	0.45 g a.i./ha (1 % label rate)	0.045 g a.i./ha (1 % label rate)	10.00
	Bidens cernua	Metsulfuron methyl	Flower bud	0.045 g a.i./ha (1 % label rate)	0.45 g a.i./ha (10 % label rate)	0.10
			Onset of flowering	0.45 g a.i./ha (10 % label rate)	0.45 g a.i./ha (10 % label rate)	1.00
	Phaseolus vulgaris	Metsulfuron methyl	Flower bud	0.45 g a.i./ha (1 % label rate)	0.045 g a.i./ha (1 % label rate)	10.00
			Onset of flowering	0.45 g a.i./ha (1 % label rate)	0.45 g a.i./ha (1 % label rate)	1.00
	Sinapis arvensis	Metsulfuron methyl	Flower bud	0.045 g a.i./ha (1 % label rate)	0.045 g a.i./ha (1 % label rate)	1.00
			Onset of flowering	0.045 g a.i./ha (1 % label rate)	0.045 g a.i./ha (1 % label rate)	1.00
	Echinochloa crus-galli	Metsulfuron methyl	Flower bud	> 0.45 g a.i./ha (1 % label rate)	> 0.45 g a.i./ha (1 % label rate)	>1.00
			Onset of flowering	0.45 g a.i./ha (1 % label rate)	> 0.45 g a.i./ha (1 % label rate)	>1.00

Appendix C. Suggested presentation of screening data

Data should be provided on all the plants routinely tested during the screening process. A list of 63 terrestrial species, 10 aquatic species and 18 forest species from 25 families tested during product development was provided (Table 1). Species selection should preferably include terrestrial weeds, crops and if possible emergent and aquatic macrophytes occurring in North America and Europe or elsewhere.

Testing of species is typically done with four to six doses in a geometric progression. Where possible, a dose–response curve should be provided along with the ER_{10}, ER_{25}, ER_{50}, slope and confidence intervals, for each species; for those species not responding in a dose–response manner, the NOER should be reported (or effect at maximum label rate).

Raw data should be provided in an electronic Excel spreadsheet. A list of species tested with common and Latin names should be provided.

Documentation should be provided on the testing procedure:

a) application method (pre-, post-emergence);
b) test substance (technical, formulated, solvent or adjuvant used);
c) indoor versus outdoor trials;
d) doses tested;
e) number of replicates per dose;
f) number of plants per dose (number of plants per pot);
g) plant growth stage at time of exposure;
h) endpoints used (definition of rating scales, quantitative or qualitative, precision);
i) seed source, stage in the plant life cycle at application (seed, seedling, leaf stages, etc.), and at recording of effect;
j) date and duration of testing, location, soil type;
k) bottom versus top watering and frequency of watering;
l) any other information pertinent to the evaluation.

C.1. Analysis of the data and triggers

Data are usually submitted as herbicidal ratings, e.g. 1 to 9, 0 to 9, 0 to 10 or 0 to 100. Scales are based upon visual observation of plant biomass, vigour, malformation, chlorosis and overall plant appearance compared with control. Herbicidal rating is converted into percentages, setting each rating to the middle of its range, as defined by registrants (see also Frans and Talbert, 1977; Boutin et al. 1993); for example:

9 defined as 100 % control = 100 % control;

8 defined as 91–99 % control = 95 % control;

7 defined as 80–90 % control = 85 % control;

6 defined as 65–79 % control = 72 % control;

5 defined as 45–64 % control = 54.5 % control;

4 defined as 30–44 % control = 37 % control;

3 defined as 16–29 % control = 22.5 % control;

2 defined as 6–15 % control = 10.5 % control;

1 defined as 1–5 % control = 3 % control;

0 defined as 0 % control = 0 % control.

Normally, each control unit will show full growth and vigour and thus will have a rating of 9 and a percentage value of 100 %. If the value is less than 100 % for any of the control units, this should be clearly stated in the results, and the possible reasons for this should be specified. It may be necessary to repeat the experiment in this case.

Statistical analysis (preferably using non-linear regression or a non-parametric analysis, e.g. Norberg-King, 1993) is performed with the data to obtain a dose–response curve for each species, namely ER_{25} and ER_{50}. Only species which are tested with four doses or more are analysed. In the results, herbicidally effective application rates (ECs) are presented only within the response range observed for the species.

Appendix D. Drift model

This appendix describes the current procedure for estimating spray drift deposition. For estimating spray drift into surface waters, individual regression curves were developed by the FOCUS Surface Water Working Group (FOCUS, 2001) for each crop grouping, as well as for each number of applications, based on fitting the various percentile drift results as a function of distance from the edge of the treated crop. Each dataset was described using a simple power function in order to obtain two regression parameters:

$$\text{Per cent drift} = A \times z^{B} \quad \text{(D.1)}$$

where per cent drift = percentile drift value (per cent of application rate) at distance z (m) from the edge of the treated field, A = regression factor (constant) and B = regression factor (exponent).

According to FOCUS (2001) this function worked well for the datasets for arable crops, vegetables (< 50 cm), vegetables (> 50 cm) and grapes (both early and late). However, a single power function with only two regression parameters seemed to be inadequate to describe the datasets for hops and fruit crops (early and late) as well as aerial applications. To represent the drift data for these cases, a regression function was developed by FOCUS (2001) using two sequential power functions splined together at a distance H:

$$\text{Per cent drift} = A \times z^{B} \text{ (for } z = 0 \text{ to } H\text{)}$$

$$= C \times z^{D} \text{ (for } z > H\text{)} \quad \text{(D.2)}$$

where per cent drift = percentage drift value (per cent of application rate) at distance z from the edge of the treated field, A = constant regression factor for distance 0 to H

B = exponential regression factor for distance 0 to H, C = constant regression factor for distance H and higher, D = exponential regression factor for distance H and higher and H = distance limit for each part of the regression (m), also called the hinge point.

This regression curve uses the regression parameters A and B to calculate drift for distances between 0 and H; regression parameters C and D are used for drift calculations for distances for H and higher. Using this approach, all of the drift datasets could be simply and accurately described by using either two parameters (arable crops, vegetables, grapes) or four parameters (hops, fruit crops and aerial application).

Table D1: Model parameters (A, B, C and D) and hinge distance (m) (FOCUS, 2001)

Crop grouping	A	B	C	D	Hinge distance (m) [(a)]
Arable and vegetable crops < 50 cm	2.7593	–0.9778	–	–	–
Hops	58.247	–1.0042	8 654.9	–2.8354	15.3
Vines, late applications and vegetables > 50 cm	44.769	–1.5643	–	–	–
Vines, early applications	15.793	–1.6080	–	–	–
Pome/stone fruit, late applications	60.396	–1.2249	210.70	–1.7599	10.3

Crop grouping	A	B	C	D	Hinge distance (m) [a]
Pome/stone fruit, early applications	66.702	–0.7520	3 867.9	–2.4183	11.4
Aerial application	50.470	–0.3819	281.1	–0.9989	16.2

(a): When a hinge distance is listed, two regression curves have been fitted to the data. The first drift regression curve uses parameters A and B and extends from the edge of the treated field to the hinge distance. The second regression curve uses parameters C and D and extends from the hinge distance to distances greater than the hinge distance.

Drift deposits for a couple of crops that have been calculated using Table D1 are presented in Table D2.

Table D2: Step 1 drift input into non-target areas with standard nozzles according to FOCUS (2001)

Crop	Distance (m) [a]	Drift (%) [b]
Pome/stone fruit, early applications	3	29.2
Pome/stone fruit, late applications	3	15.7
Field crops	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 > 5 cm)	3	8.0

(a): Distance from edge of crop to non-target area.
(b): Percentage of the application dose.

The equations presented above (equations D.1 and D.2) do not directly allow the calculation of buffer strips necessary to meet maximum acceptable loads. However, that is possible when using equation D.3, which can be obtained from the previous equations after transformation.

$$d = \left(\frac{Maxload \cdot 10000}{Appdose \cdot D_{\mathrm{Re}d} \cdot A} \right)^{\frac{1}{B}} \qquad \text{(D.3)}$$

where

d: necessary width of buffer strip (m);

Maxload: maximum acceptable deposition (kg/ha);

Appdose: application dose (kg/ha);

A: regression factor (constant);

B: regression factor (exponent);

D_{red}: nozzle drift reduction (%).

If the regression curve includes a hinge point as shown in equation D.2 the calculation in equation D.3 has to be performed twice considering the regression factors A and B as well as factors C and D. After the calculation has been performed, the calculated distance that falls into the respective range of the regression curve should be used.

Table D3 shows examples for different crops when 1 kg/ha was sprayed using nozzles with 75 % drift reduction and the maximum acceptable load was calculated to be 0.01 kg/ha.

Table D3: Example for necessary drift buffers (m) for different crops (75 % nozzle drift reduction) [(a)]

Crop type	Necessary distance (m)	Hinge distance(m)	Calculated with regression constants
Field	0.7	–	A,B
Hop	14.4	15.3	A,B
Orchard (early stage)	17.2	11.4	C,D
Orchard (late stage)	9.2	10.3	A,B

(a): Application dose 1 kg/ha, maximum acceptable load: 0.01 kg/ha.

Appendix E. Considering multiple applications in the risk assessment

E.1. Introduction

Multiple applications of a PPP will also lead to multiple exposure events in off-field areas owing to drift. In line with the general approaches for other environmental compartments and organism groups, this also needs to be adequately considered in the risk assessment for non-target terrestrial plants. As no specific guidance on that issue was provided in previous EU guidance (Guidance Document on Terrestrial Ecotoxicology under Council Directive 91/414/EEC, SANCO/10329/2002 rev. 2 final), possible options for addressing the risk owing to multiple exposure events are discussed in this appendix.

The most straightforward way to assess the effects of multiple exposure to a compound might be by directly incorporating the exposure scheme into the design of the toxicity tests, as proposed in the current OECD test guidelines 208 and 227 for limit testing. There are, however, several aspects that would advise against using the results from such multiple-exposure tests in a lower tier risk assessment for terrestrial plants. Vegetative vigour tests conducted in accordance with the relevant test guideline, OECD 227, are performed with young plants grown from seed usually to the two to four true leaf stage, which should thus be considered a sensitive reference stage. Owing to the growth of test plants, additional exposure events would therefore have an impact on later growth stages, making extrapolation of the results to other plants difficult. While the recommended plant density for plants at the two to four true leaf stage is optimised in standard test protocols to minimise shading, such consistent worst-case exposure conditions can no longer be ensured for plants at a later growth stage. In addition, the specific growth stage and growth rate of test plants at the time of application will have an impact on the results, thus increasing their overall uncertainty. Such issues could only be properly addressed in a higher tier assessment.

It is therefore necessary, at least for the lower tier risk assessment, to develop approaches that make use of the toxicity information from standard testing (i.e. with a single application at a sensitive reference growth stage) and address the impact from multiple exposure events via appropriate extrapolation factors (e.g. expressed as multiple application factors, MAFs). There are basically two approaches for extrapolating, namely:

- an approach referring to application rates and environmental fate parameters of a compound, taking into account residue kinetics to calculate an overall exposure level that would determine an overall level of effects (fate-based approach, using a multiple-application factor, MAF-fate); and
- an approach referring to effect levels caused by individual exposure events, assuming direct cumulation of those effect levels, whereby the recalculation of exposure to effect levels and vice versa is based on the non-linear rate–response curve (effect-based approach, using a multiple-application factor, MAF-effect).

The applicability of both approaches for different areas of the NTTP risk assessment is discussed below, and guidance on the actual implementation of each approach in its relevant context is provided.

The effect-based approach and the application of the concept for assessments based on aggregated data or on results from limit testing, as described under the following points, were developed by a group of authors including Klaus Swarowsky, Andreas Höllrigl-Rosta, Carmen Schweikert (German Federal Environment Agency, UBA) and Peter Craig (University of Durham, Department of Mathematical Sciences); please refer to Swarowsky et al. (2014a,b) and Höllrigl-Rosta et al. (2014). The written draft was provided to the Working Group on NTTP Risk Assessment, which adapted it by their means for the NTTP risk assessment.

E.2. Fate-based vs. effect-based approach

The fate-based approach for assessing the risk due to multiple applications of a PPP is already established in various areas of environmental risk assessment, i.e. when characterising the risk for non-target arthropods, aquatic organisms, birds and mammals. It basically assumes that the effect level owing to a series of exposure events is similar to (or covered by) the effect level owing to a single exposure event at a higher level (calculated with or without consideration of in-between degradation). Such an approach is useful for situations in which there is only a limited potential for a carry-over of effects between the individual exposure events, e.g. when different parts of a local population are exposed at different levels during each event, or when rapid recovery from temporary effects can be assumed between the events, which could be the case for very low effect levels caused by a single exposure level or if only a specific relatively short developmental stage is affected.

It is also a meaningful approach when a risk assessment is based on a toxicity value from a test with continuous exposure (such as dietary tests with birds or mammals or flow-through tests with aquatic organisms). In such a case, no critical effects would be expected from a series of exposure events in the field, as long as the highest predicted exposure level stays below the relevant (no) effect level from the test.

Summing up, the fate-based approach may be applied where it can be reasonably assumed that the risk for non-target organisms exposed to a compound via environmental media such as water, soil or food is actually determined by the maximum predicted exposure level (i.e. a single peak exposure level considered to represent the whole actual exposure pattern) of that compound in those media. For the NTTP risk assessment, such a case would be the evaluation of effects on germination and growth based on seedling emergence tests conducted in accordance with OECD 208. Due to the short duration of the germinating phase in relation to typical vertical leaching rates and application intervals for herbicides, it can be reasonably assumed that germinating seedlings in soil will normally not encounter multiple exposure events.

However, a different situation has to be considered where a carry-over of effects between individual exposure events cannot be excluded and/or where the environmental fate of a test compound is already integrated in the result from a toxicity test. This is in particular the case for effects on the growth of terrestrial plants, as evaluated on the basis of vegetative vigour tests conducted in accordance with OECD 227. In those tests, the effect levels owing to a single spray application are determined only once at the end of the test after 21 to 28 days, thus integrating all possible translocation, damage, and recovery processes in the test plants during that period. No conclusion can thus be drawn on the time course of effect levels between initial application and effect evaluation at the end of the study. Furthermore, sprayed plants are not exposed via an external medium but directly over their leaf surfaces, which means that the effect values from the study are determined by the fate of the compound on the plants as well as by its toxicokinetics within them.

For a risk assessment, it must therefore be assumed that the endpoint from a vegetative vigour test directly corresponds to the initial exposure level in that test, integrating all possible fate and toxicokinetic processes during the test period. This also includes, in particular, recovery processes within the observation period. Therefore, it must be assumed that a sequence of exposure events at an interval that is shorter than 21 to 28 days will result in a cumulation of effects in such a way that the effect due to each exposure event is calculated as a percentage referring to the already affected status from previous exposure events. Such a cumulation of effects is calculated based on the assumption that the sensitivity of the plant does not increase over multiple applications, i.e. the effect level (biomass reduction at test termination) per unit application rate remains constant.

Like the fate-based approach, the effect-based approach also includes some simplifications and would benefit from further scientific input. Nevertheless, in the case of the NTTP vegetative vigour test, the effect-based MAF makes better use of the available data and is thus deemed scientifically better justified than the fate-based MAF.

It is therefore recommended that a fate-based approach is used for assessing multiple applications on the basis of seedling emergence tests conducted in accordance with OECD 208, whereas an effect-based approach should be applied for assessing multiple applications on the basis of vegetative vigour tests according to OECD 227. Still, it appears advisable to check the appropriateness of any MAF concept for NTTP risk assessment using appropriate reference data, i.e. standard laboratory tests with multiple applications. Such analysis could also try to elucidate actual time courses of effects on plants after the exposure peak(s), in order to achieve a better understanding of the parameters driving the overall effect.

E.3. Fate-based methods for assessing the risk from multiple applications

As discussed above, the fate-based approach is recommended for assessing multiple applications of a PPP on the basis of seedling emergence tests according to OECD 208. The ER_x values from that test are typically expressed as deposition rates in g/ha or kg/ha. In order to obtain an estimate for an off-field deposition rate that reflects cumulation of residues in off-field soil due to multiple applications, the apparently most straightforward approach would be to make use of the calculated PEC_{soil} values from the environmental fate section in combination with applicable drift percentages.

For PEC_{soil} calculation, please refer to the respective documents from EFSA (e.g. EFSA, 2012a)

E.4. Effect-based method for assessing the risk from multiple applications

E.4.1. Conceptual basis of the approach

E.4.1.1. Probit model

Other than the fate-based approach, the effect-based approach does not take into account the time interval between two applications. It is built on the assumption that recovery from effects is not relevant on the time scale of vegetative vigour tests conducted according to OECD 227, with an observation interval of 21 to 28 days between exposure and effect evaluation. Consequently, it is assumed for assessing the possible cumulation of effects due to multiple applications of a PPP at intervals less than 21 to 28 days that the magnitude of effects at the time of a subsequent application is equal to the effect level as observed in the test at the end of the observation period.

The basic concept of the effect-based approach for assessing the risk for organisms due to multiple exposure to a compound is linking exposure and effect levels via the dose–response curve derived from toxicological testing. It is assumed that this curve is appropriately described by the probit model, in which the dependency of the effect level p from exposure level c is defined by the distribution function Φ of a normal (0,1) distribution, the parameters a (slope) and b (intercept) determining the shape and the central position of the distribution, respectively.

$$p = \Phi(a \ln c + b) \qquad \Leftrightarrow \qquad \Phi^{-1}(p) = a \ln c + b$$

It is thus possible to derive the slope and intercept parameters from two points (ER_x and ER_y, representing the exposure levels leading to x % and y % effect, respectively) on the rate–response curve, according to the following equations:

$$\Phi^{-1}(x) = a \ln ER_x + b \quad \Leftrightarrow \quad \ln ER_x = \frac{\Phi^{-1}(x) - b}{a}$$

$$\ln \frac{ER_x}{ER_y} = \frac{\Phi^{-1}(x) - \Phi^{-1}(y)}{a} \quad \Leftrightarrow \quad a = \frac{\Phi^{-1}(x) - \Phi^{-1}(y)}{\ln ER_x - \ln ER_y}$$

$$b = \Phi^{-1}(x) - a \ln ER_x$$

As not only ER_{50} but also ER_{25} values are routinely available from vegetative vigour tests conducted in accordance with OECD 227, slope and intercept of the probit model can be easily calculated for each tested plant species.

E.4.1.2. Logit model

It appears, in principle, advisable that the same rate–response model that was used for the derivation of ER_x values from a test is also used for estimating effect cumulation due to multiple exposure. The basic equations for the logit model can be easily derived from those for the probit model by substituting the general term $\Phi^{-1}(x)$ by the general term $\ln(x/1-x)$. The basic logit function is thus described by:

$$p = \frac{1}{1+e^{-(a\ln c+b)}} \quad \Leftrightarrow \quad \ln\frac{p}{1-p} = a\ln c + b$$

Slope a and intercept b are calculated as follows from available ER_x values:

$$\ln\frac{x}{1-x} = a\ln \mathrm{ER}_x + b \quad \Leftrightarrow \quad \ln \mathrm{ER}_x = \frac{\ln\frac{x}{1-x} - b}{a}$$

$$\ln\frac{\mathrm{ER}_x}{\mathrm{ER}_y} = \frac{\ln\frac{x}{1-x} - \ln\frac{y}{1-y}}{a} \quad \Leftrightarrow \quad a = \frac{\ln\frac{x}{1-x} - \ln\frac{y}{1-y}}{\ln \mathrm{ER}_x - \ln \mathrm{ER}_y}$$

$$b = \ln\frac{x}{1-x} - a\ln \mathrm{ER}_x$$

It may also be the case that actual ER_{25} and ER_{50} values were estimated using non-linear regression models other than probit or logit. Nevertheless, it is considered that the sigmoidal shape of any such regression curve is sufficiently approximated within the effect range relevant for risk management decisions (*c*. 10–50 %) by a probit or logit curve that has been fitted to those given ER_{25} and ER_{50} estimates.

E.4.2. Recalculations between exposure and effect levels and cumulation of effects

To estimate cumulated effects on NTTPs due to multiple exposure to a compound, it is first necessary to separately calculate the effect levels for each individual exposure event. To this end, each single application rate (AR_i) is first multiplied with the appropriate drift percentage (DP), taking into account the number of applications (n). Then, the single effect levels can be calculated using the previously derived slope (a) and intercept (b) data from the rate–response curve:

$$p_i = \Phi[a\ln(\mathrm{AR}_i \times \mathrm{DP}) + b] \qquad \text{(probit model)}$$

$$p_i = \frac{1}{1+e^{-(a\ln(\mathrm{AR}_i\times\mathrm{DP})+b)}} \qquad \text{(logit model)}$$

For estimating cumulative effect levels, reference is made to the well-established toxicological Independent Action (IA) concept (see Kortenkamp et al., 2009), which is based on the two core assumptions that the toxicity of each of a number of simultaneously acting compounds is not influenced by the presence of the other compounds, and that all compounds affect the same biological endpoint. The cumulative effect level (p_{cum}) according to the IA concept for n compounds causing individual effect levels p_i is then calculated as follows:

$$p_{cum} = 1 - \prod_{n} (1 - p_i)$$

As discussed above, the effect-based approach is built on the assumption that recovery from effects is not relevant on the timescale of vegetative vigour tests conducted in accordance with OECD 227, while the sensitivity of plants to subsequent exposure events is not affected by previous exposure. Hence, repeated exposure to the same compound on this timescale can be considered equivalent to simultaneous exposure to several different compounds. The second provision that these several different compounds must affect the same biological endpoint to apply the IA approach is always fulfilled in the case of subsequent multiple exposure to the same compound. It is therefore concluded that the IA approach can also be applied for estimating cumulative effect levels due to subsequent exposure to the same compound. For a series of *n* equal drift values, the single effect levels *p* are equal and are combined into a cumulative effect (p_{cum}) according to the following simplified IA formula:

$$p_{cum} = 1 - (1 - p)^n$$

Finally, the cumulative effect level is recalculated to a corresponding cumulative effect rate (CR), which will be used for assessing the acceptability of exposure levels.

$$\ln \text{ CR} = \frac{\Phi^{-1}(p_{cum}) - b}{a} \quad \Leftrightarrow \quad \text{CR} = \exp\left(\frac{\Phi^{-1}(p_{cum}) - b}{a}\right) \qquad \text{(probit model)}$$

$$\ln \text{ CR} = \frac{\ln \frac{p_{cum}}{1 - p_{cum}} - b}{a} \quad \Leftrightarrow \quad \text{CR} = \exp\left(\frac{\ln \frac{p_{cum}}{1 - p_{cum}} - b}{a}\right) \qquad \text{(logit model)}$$

As an example the twofold application of 100 g/ha in an arable crop of a compound with an ER_{50} of 50 g/ha and an ER_{25} of 15 g/ha is assessed as follows, applying the probit model. First, slope (*a*) and intercept (*b*) are calculated from the two ER_x values:

$$a = \frac{\Phi^{-1}(x) - \Phi^{-1}(y)}{\ln \text{ ER}_x - \ln \text{ ER}_y} = \frac{\Phi^{-1}(0.5) - \Phi^{-1}(0.25)}{\ln \text{ 50} - \ln \text{ 15}} = 0.560$$

$$b = \Phi^{-1}(x) - a \ln \text{ ER}_x = \Phi^{-1}(0.5) - 0.560 \times \ln \text{ 50} = -2.192$$

Second, the drift rate per application is determined as 2.38 g/ha, using the appropriate drift percentile for the twofold application, and the corresponding effect percentage is calculated:

$p_i = \Phi[0.560 \times \ln (100 \times 0.0238) - 2.192] = 4.4\ \%$

Two successive effect levels of 4.4 % will lead to a cumulative effect of $p_{cum} = 1 - (1 - 0.044)^2 = 0.086$, which is then recalculated to a (virtual) cumulative rate as follows:

$$\text{CR} = \exp\left(\frac{\Phi^{-1}(0.086) + 2.192}{0.560}\right) = 4.4\,\text{g/ha}$$

Notably, this recalculated cumulative rate is slightly (by a factor of 1.2) higher than a PER based on the same application rate and taking into account an application interval of 14 days and a DT_{50} of the compound in soil of 43.5 days (see above).

E.4.3. Derivation of appropriate rate–response parameters for larger datasets

It is one of the characteristics of the NTTP risk assessment that it is usually based on the toxicity data of more than one species, to account for the high diversity of the plant realm. One of the preferred approaches for applying toxicity data for several plant species is the calculation of SSDs from the available ER_x values, from which HC_5 values can be derived that cover the sensitivity of 95 % of all species with regard to that endpoint. In other words, only 5 % of all species are assumed to have an ER_x that is equal to or smaller than the HC_5. However, information on the rate–response behaviour of individual species, which are required for performing the calculations on effect cumulation, are not visible from a single HC_5 figure. It is therefore necessary to develop approaches for estimating rate–response parameters that can be used in combination with HC_5 values for assessing effects due to multiple exposure.

Major challenges in this respect are the non-linear interdependencies of the parameters determining the cumulation of effects due to a series of exposure events. As obvious from the equations above, probit/logit slope and intercept values are correlated via the ER_{50}, and the ratio of the recalculated (virtual) cumulative exposure rate to the single exposure rate depends on the single-exposure effect level and the shape of the rate–response curve. As a consequence, no feasible generic approach could be developed for deriving appropriate rate–response parameters to be used in connection with HC_5 effect values.

Instead, a pragmatic approach whereby the rate–response data are derived from two HC_5 values for the 25 % (HC_5-ER_{25}) and the 50 % effect level (HC_5-ER_{50}) was evaluated. Basically, this approach assumes that the individual sensitivities and rate–response relationships of the tested plant species are represented by the sensitivity and rate–response relationship of a virtual HC_5 species. While the sensitivity of that virtual HC_5 species, by definition, reflects the 5th percentile of the assumed sensitivity distribution, no such statement is possible for the rate–response relationship, because the individual rate–response relationships of the tested plants may vary independently of their respective sensitivity. Hence, the applicability of the approach was assessed empirically, using a dataset with ER_{25} and ER_{50} values from vegetative vigour tests with 23 herbicidal products containing active substances from 10 mode-of-action classes (Table E1).

Table E1: Modes of action of herbicides used in the analysis of rate–response parameters

Mode of action [(a)]	No of products
Acetolactate synthase (ALS) or acetohydroxy acid synthase (AHAS) inhibitor	6
Photosystem II inhibitor	3
Photosystem II inhibitor	4
Synthetic auxin	3
Protoporphyrinogen oxidase (PPG oxidase or protox) inhibitor	2
Acetyl CoA carboxylase (ACCase) inhibitor	1
ALS or AHAS inhibitor + mitosis inhibitor	1
Carotenoid biosynthesis inhibitor + mitosis inhibitor	1
Enolpyruvyl shikimate-3-phosphate (EPSP) synthase inhibitor	1
Fatty acid and lipid biosynthesis inhibitor	1
Mitosis inhibitor	1

(a): Classification of herbicide mode of action according to the Weed Science Society of America classification scheme (WSSA, 2011 http://wssa.net/wp-content/uploads/WSSA-Mechanism-of-Action.pdf.)

The analysis was performed by estimating HC_5-ER_{25} and HC_5-ER_{50} values from the available true ER_{25} and ER_{50} values for tested plant species. Single tests yielding greater than figures for ER_{50} or both endpoints were disregarded, even if that resulted in a lower number of data points for deriving the SSD; this was considered acceptable, as the focus of the analysis was not on generating HC_5 values for quantitative risk assessments but on producing ratios of HC_5-ER_{25} and HC_5-ER_{50} values. In other words, the SSD approach was used as a method for aggregating the given toxicity datasets.

Uncertainties with regard to the representativeness of HC_5 figures for the entirety of non-tested species would thus not be critical in this context. Probit and logit slope and intercept values, respectively, were calculated from the ER_{25} and ER_{50} of each tested plant species as well as from the HC_5-ER_{25} and the HC_5-ER_{50}. To obtain a basis for comparison, a scenario was defined with three exposure events at an exposure level per event that would correspond to a 10 % effect level for the virtual HC_5 species (i.e. the HC_5-ER_{10}). Individual effect levels per exposure event and the cumulated effect levels after three exposure events were calculated for each set of rate–response data for tested plants and virtual HC_5 species, respectively. The cumulated effect levels were selected as reference endpoints for comparison, because they also constitute the relevant endpoint for risk management decisions. Other parameters and calculation results were considered not suitable for comparison, owing to their various non-linear interdependencies, as mentioned before. A summary of the results is provided in Table E2.

Table E2: Comparison of calculated cumulated effects from threefold exposure to the HC_5-ER_{10}, considering rate–response parameters (rrps) derived from HC5-ER_{25} and HC_5-ER_{50} versus rrps for individual tested plant species

Herbicide type (mode of action)	Cumulated effects from threefold exposure to HC_5-ER_{10} (%)				No of species considered
	rrp from HC_5 data	rrp from individual species data			
		Median	90th percentile	95th percentile	
Photosystem II inhibitor	27.1	2.3	16.9	21.1	4
Synthetic auxins	27.1	5.5	19.1	23.0	5
ALS or AHAS inhibitors	27.1	5.1	11.6	12.3	4
ALS or AHAS inhibitors	27.1	8.6	13.4	13.7	6
ALS or AHAS inhibitors	27.1	0.6	18.5	21.5	7
ALS or AHAS inhibitors	27.1	0.1	12.3	14.1	6
Enolpyruvyl shikimate-3-phosphate (EPSP) synthase inhibitors	27.1	2.8	17.2	17.5	11
Photosystem II inhibitor	27.1	1.6	17.1	19.0	6
Synthetic auxins	27.1	4.6	14.4	15.6	3
Photosystem II inhibitor	27.1	0.0	2.7	3.9	7
Fatty acid and lipid biosynthesis inhibitors	27.1	0.4	1.2	1.3	3
Carotenoid biosynthesis inhibitors + mitosis inhibitors	27.1	24.8	25.2	25.2	3
Photosystem I inhibitors	27.1	0.0	13.3	18.2	6
ALS inhibitors mitosis inhibitors	27.1	6.9	23.9	**33.8**	9
ALS or AHAS inhibitors	27.1	0.4	11.5	13.4	5
Photosystem II inhibitor	27.1	0.0	4.5	4.7	6
Synthetic auxins	27.1	3.0	10.3	11.6	9
ALS or AHAS inhibitors	27.1	7.6	**30.5**	**31.3**	7
Mitosis inhibitors	27.1	0.0	17.5	24.2	6
ALS or AHAS inhibitors	27.1	16.5	20.1	21.2	5
Protoporphyrinogen oxidase (PPG oxidase or protox) inhibitors	27.1	2.2	14.0	15.5	3
Acetyl CoA carboxylase (ACCase) inhibitors	27.1	1.8	4.1	5.0	6

Protoporphyrinogen oxidase (PPG oxidase or protox) inhibitors	27.1	0.0	0.5	0.8	6

The cumulated effect levels in the scenario with three exposure events at the HC_5-ER_{10}, as calculated with the rate–response data for the virtual HC_5 species are in all 23 cases markedly greater than the geometric mean of the cumulated effect levels calculated with the rate–response data for the individual tested plant species. In 21 out of 23 cases, the results for the virtual HC_5 species are also greater than the 95th percentile of the cumulated effect levels for the individual tested plant species. There are only two cases in which they are below the 95th percentile and in one case slightly below the 90th percentile for the individual plant species (marked in bold in the table above). Closer inspection of those cases reveals that this can be ascribed to the occurrence of one plant species with a low ER_{25} and a relatively high ER_{50} in the respective dataset (heterogeneous dataset). However, the actual cumulated effect levels up to 33.8 % (probit) or 34.2 % (logit) do not differ much from the reference level of 27.1 % for three consecutive effects at the 10 % level and would not constitute a significant difference in terms of regulatory consequences.

It can also be stated that the deviation between the reference level of 27.1 % and the 95th percentile of cumulated effect levels for the individual tested plant species (or margin of safety, MoS) is relatively low (≤ 25 %) in 8 out of 23 cases. A larger deviation (> 25 % and ≤ 50 % for MoS) is observed in a further eight cases. There are only 5 out of 23 cases in which this deviation is ≥ 80 %, which might be considered to reflect a relevant overestimation of effect cumulation. For such cases, it may be a meaningful refinement approach to identify the highest actual potential for effect cumulation at a given exposure level by considering the rate–response data for each of the tested species for which such data can be generated (note that this will require performing the calculations for all of those species, because the worst case in terms of effect cumulation cannot be easily predicted from the ER_{25} and ER_{50} values alone).

Summing up, according to our analysis, the 95th percentile of cumulated effect levels for the individual tested plant species could be replaced by a surrogate value, based on rate–response data for a "virtual HC_5 species" in a realistic worst-case scenario for herbicide applications. A sufficient MoS exists for homogeneous toxicity datasets with similar rate–response curves for individual tested species. However, the margin of safety can become much smaller or the approach might even become unsuitable for heterogeneous datasets in combination with critical exposure scenarios. Additional research is thus required for a better definition of the possible limitations of the approach.

E.4.4. Application of the effect-based approach—relevant rates for limit tests

To reduce the effort required to assess the risk to NTTPs for compounds with relatively low toxicity, screening tests at field rates may be applied. However, owing to the necessity of considering an assessment or safety factor and also possible effect cumulation from multiple exposures, it cannot be generally concluded that a screening test with ≤ 50 % effect at field rates would sufficiently cover the risk for NTTPs from drift deposits in off-field areas. An approach is described in this section to estimate the required application rate (AR) in a limit test that would allow conclusions to be drawn on an acceptable risk for NTTPs in off-field areas.

For *n* applications of a PPP, provided that each AR is identical, the single effect levels (*y*) adding up to the acceptable overall effect level (cumulatively) can be calculated as follows, based on the assumptions of the effect-based approach:

$$y = 1 - \sqrt[n]{1 - [acceptable \quad effect \quad level]}$$

Assuming a limit test that results in an effect level of *x*, an acceptable limit test rate, LTR_x, can be defined, for which the effect level at the field rate (AR) will not exceed the acceptable effect level:

$$LTR_x = AR \times \frac{ER_x}{ER_y}$$

Applying the probit model, algebraic transformation of the term ER_x/ER_y (see above) results in:

$$LTR_x = AR \times \exp\left(\frac{\Phi^{-1}(x) - \Phi^{-1}(y)}{a}\right) \quad \text{where } y = 1 - \sqrt[n]{1 - [acceptable \quad effect \quad level]}$$

$$LTR_x = AR \times \exp\left(\frac{\Phi^{-1}(x) - \Phi^{-1}(1 - \sqrt[n]{1 - [acceptable \quad effect \quad level]})}{a}\right)$$

As pointed out above, the assessment of the risk for off-field plants must consider an assessment of the safety factor (SF) and is then performed for the deposited drift percentages (DPs) in place of the field application rate (AR). Hence, the latter is substituted in the formula by the term (AR × DP × SF):

$$LTR_x = AR \times DP \times SF \times \exp\left(\frac{\Phi^{-1}(x) - \Phi^{-1}(1 - \sqrt[n]{1 - [acceptable \quad effect \quad level]})}{a}\right)$$

Making use of the assumptions of the effect-based approach, this formula can thus be used to calculate the limit test rate, for which an observed effect level ≤ *x* would ensure that the acceptable effect level for NTTPs in off-field areas is not breached. As described above, the corresponding equations for applying the logit model (not shown here) can be derived by substituting the general term $\Phi^{-1}(x)$ in the probit equations by the general term $\ln(x/1-x)$.

For the slope parameter, a default estimate was derived from the same dataset as that used for the validation exercise presented above. To do so, the empirical 95th percentile of slopes was derived from all rate–response relationships in the datasets, leading to a probit slope of 0.3775, a logit slope of 0.6148, or a ratio ER_{50}/ER_{10} of around 30.

As an example, the concept is applied to an intended use with *n* = 3 applications of a PPP in an arable crop, using the above derived slope. Thus, taking into account a probit slope of 0.3775 and an acceptable effect level of 50 %, a variable safety factor, and a drift percentile of 2.01 % for three applications, and assuming that an effect level of 50 % was observed in the limit test, the following result is obtained:

$$LTR_{0.5} = AR \times 2.01\% \times SF \times \exp\left(\frac{\Phi^{-1}(0.5) - \Phi^{-1}(1 - \sqrt[3]{1 - [0.5]})}{0.3775}\right)$$

$$LTR_{0.5} = AR \times SF \times 0.18$$

In other words, a limit test conducted at a rate of *c*. 18 % of the field application rate times the safety factor considered in the risk assessment, and resulting in an effect level of ≤ 50 %, would ensure that the acceptable effect level is not breached for NTTPs in the off-field area.

While it would theoretically be possible to perform such calculations for any effect level observed in a limit test, it must be kept in mind that quantifying a low magnitude of effects ≤ 25 % with sufficient certainty could be problematic with current test methods. It is, therefore, proposed only to consider effect level classes ≤ 25 % and ≤ 50 % in limit tests at field rates for this type of calculation. If it can

be demonstrated with sufficient certainty that no effect at all (no growth impairment, no signs of phytotoxicity) would occur at the limit test rate, a numerical level of ≤ 10 % can be considered.

The following tables compile the multiplication factors that must be applied to field application rates in order draw conclusions on an acceptable risk for NTTPs in the off-field area, provided that the respective limit test at the rate of multiplication factor ×AR results in an effect level of ≤ 50 %, ≤ 25 % or ≤ 10 % (representing a proxy for no visible effect at all), respectively. Generic probit and logit slopes $a_{probit} = 0.3775$ and $a_{logit} = 0.6148$, corresponding to a ratio of $ER_{50}/ER_{10} = 30$ were used in the calculations.

Table E3: Proportion of field rate to be covered by a limit test to ensure ≤ 50 % overall effect level for different spray drift scenarios when the field rate = ER_{10}, probit model (figures must be multiplied by the appropriate safety factor)

Number of applications	1	2	3	4	5	6	7	8	9	10	11	12
Arable crops	0.01	0.03	0.05	0.07	0.10	0.12	0.15	0.17	0.20	0.23	0.26	0.29
Arable crops × 2	0.02	0.06	0.10	0.15	0.19	0.24	0.29	0.33	0.39	0.45	0.51	0.57
Arable crops > 900 L/ha	0.01	0.04	0.08	0.09	0.12	0.14	0.18	0.18	0.21	0.24	0.28	0.31
Vines	0.02	0.09	0.17	0.26	0.37	0.47	0.58	0.69	0.81	0.93	1.05	1.18
Orchards, early	0.08	0.30	0.59	0.93	1.28	1.66	2.06	2.44	2.86	3.30	3.74	4.18
Orchards, late	0.04	0.14	0.27	0.40	0.54	0.67	0.83	0.95	1.12	1.28	1.45	1.63
Hops	0.05	0.21	0.39	0.60	0.84	1.09	1.33	1.48	1.74	2.01	2.27	2.54
Knapsack sprayer, plants < 50 cm	0.01	0.03	0.05	0.07	0.10	0.12	0.15	0.17	0.20	0.23	0.26	0.29
Knapsack sprayer, plants > 50 cm	0.02	0.09	0.17	0.26	0.37	0.47	0.58	0.69	0.81	0.93	1.05	1.18
Knapsack sprayer, spraying screen	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01
Gardening, plants < 50 cm	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02
Gardening, plants > 50 cm	0.00	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.04	0.04	0.05	0.05
Gardening, trees early < 2 m	0.04	0.11	0.18	0.28	0.37	0.46	0.55	0.63	0.74	0.85	0.97	1.08
Gardening, trees early > 2 m	0.11	0.33	0.59	0.87	1.19	1.49	1.82	2.15	2.53	2.91	3.30	3.69
Gardening, trees late	0.01	0.02	0.04	0.05	0.07	0.08	0.09	0.10	0.11	0.13	0.15	0.17

Table E4: Proportion of field rate to be covered by a limit test to ensure ≤ 50 % overall effect level for different spray drift scenarios when the field rate = ER_{25}, probit model (figures must be multiplied by the appropriate safety factor)

Number of applications	1	2	3	4	5	6	7	8	9	10	11	12
Arable crops	0.01	0.05	0.09	0.13	0.18	0.22	0.27	0.31	0.36	0.41	0.47	0.52
Arable crops × 2	0.03	0.10	0.18	0.27	0.36	0.44	0.54	0.61	0.72	0.83	0.94	1.05
Arable crops > 900 L/ha	0.02	0.08	0.15	0.16	0.22	0.26	0.32	0.33	0.39	0.45	0.51	0.57
Vines	0.04	0.16	0.31	0.48	0.67	0.86	1.06	1.26	1.48	1.70	1.93	2.16
Orchards, early	0.15	0.55	1.07	1.70	2.36	3.04	3.79	4.47	5.25	6.05	6.85	7.67
Orchards, late	0.08	0.26	0.49	0.73	0.99	1.23	1.52	1.74	2.05	2.36	2.67	2.99
Hops	0.10	0.38	0.71	1.11	1.54	1.99	2.44	2.72	3.20	3.68	4.17	4.67
Knapsack sprayer, plants < 50 cm	0.01	0.05	0.09	0.13	0.18	0.22	0.27	0.31	0.36	0.41	0.47	0.52
Knapsack sprayer, plants > 50 cm	0.04	0.16	0.31	0.48	0.67	0.86	1.06	1.26	1.48	1.70	1.93	2.16
Knapsack sprayer, spraying screen	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02
Gardening, plants < 50 cm	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03
Gardening, plants > 50 cm	0.00	0.01	0.02	0.03	0.04	0.04	0.05	0.06	0.07	0.08	0.09	0.10
Gardening, trees early < 2 m	0.07	0.20	0.33	0.51	0.68	0.84	1.01	1.16	1.36	1.56	1.77	1.98
Gardening, trees early > 2 m	0.19	0.60	1.07	1.59	2.18	2.73	3.34	3.95	4.64	5.34	6.05	6.77
Gardening, trees late	0.02	0.05	0.07	0.10	0.12	0.14	0.16	0.18	0.21	0.24	0.27	0.31

Table E5: Proportion of field rate to be covered by a limit test to ensure ≤ 50 % overall effect level for different spray drift scenarios when the field rate = ER_{50}, probit model (figures must be multiplied by the appropriate safety factor)

Number of applications	1	2	3	4	5	6	7	8	9	10	11	12
Arable crops	0.03	0.10	0.18	0.26	0.35	0.43	0.53	0.60	0.71	0.81	0.92	1.03
Arable crops × 2	0.06	0.20	0.35	0.52	0.70	0.86	1.06	1.20	1.41	1.62	1.84	2.06
Arable crops > 900 L/ha	0.04	0.16	0.30	0.32	0.42	0.52	0.63	0.65	0.76	0.88	0.99	1.11
Vines	0.08	0.31	0.61	0.95	1.32	1.68	2.07	2.47	2.90	3.34	3.79	4.24
Orchards, early	0.29	1.08	2.11	3.34	4.63	5.98	7.44	8.78	10.32	11.87	13.46	15.05
Orchards, late	0.16	0.52	0.97	1.43	1.95	2.42	2.98	3.42	4.02	4.62	5.24	5.86
Hops	0.19	0.75	1.40	2.17	3.03	3.91	4.79	5.34	6.28	7.22	8.19	9.16
Knapsack sprayer, plants < 50 cm	0.03	0.10	0.18	0.26	0.35	0.43	0.53	0.60	0.71	0.81	0.92	1.03
Knapsack sprayer, plants > 50 cm	0.08	0.31	0.61	0.95	1.32	1.68	2.07	2.47	2.90	3.34	3.79	4.24
Knapsack sprayer, spraying screen	0.00	0.00	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.03	0.04	0.04
Gardening, plants < 50 cm	0.00	0.01	0.01	0.02	0.02	0.03	0.04	0.04	0.05	0.05	0.06	0.07
Gardening, plants > 50 cm	0.01	0.02	0.04	0.06	0.07	0.08	0.10	0.11	0.13	0.15	0.17	0.19
Gardening, trees early < 2 m	0.14	0.39	0.65	1.00	1.34	1.65	1.98	2.27	2.67	3.07	3.48	3.89
Gardening, trees early > 2 m	0.38	1.18	2.11	3.12	4.27	5.37	6.55	7.75	9.10	10.48	11.87	13.28
Gardening, trees late	0.04	0.09	0.14	0.19	0.24	0.28	0.31	0.35	0.41	0.48	0.54	0.60

Table E6: Proportion of field rate to be covered by a limit test to ensure ≤ 50 % overall effect level for different spray drift scenarios when the field rate = ER_{10}, logit model (figures must be multiplied by the appropriate safety factor)

Number of applications	1	2	3	4	5	6	7	8	9	10	11	12
Arable crops	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.10	0.12	0.13
Arable crops × 2	0.01	0.02	0.04	0.05	0.07	0.09	0.12	0.14	0.17	0.20	0.23	0.27
Arable crops > 900 L/ha	0.00	0.02	0.03	0.03	0.05	0.06	0.07	0.08	0.09	0.11	0.13	0.14
Vines	0.01	0.03	0.06	0.10	0.14	0.18	0.23	0.29	0.35	0.41	0.48	0.55
Orchards, early	0.03	0.11	0.21	0.35	0.49	0.65	0.84	1.02	1.23	1.46	1.70	1.95
Orchards, late	0.02	0.05	0.10	0.15	0.21	0.26	0.34	0.40	0.48	0.57	0.66	0.76
Hops	0.02	0.08	0.14	0.22	0.32	0.43	0.54	0.62	0.75	0.89	1.03	1.19
Knapsack sprayer, plants < 50 cm	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.10	0.12	0.13
Knapsack sprayer, plants > 50 cm	0.01	0.03	0.06	0.10	0.14	0.18	0.23	0.29	0.35	0.41	0.48	0.55
Knapsack sprayer, spraying screen	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Gardening, plants < 50 cm	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
Gardening, plants > 50 cm	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02
Gardening, trees early < 2 m	0.02	0.04	0.07	0.10	0.14	0.18	0.22	0.26	0.32	0.38	0.44	0.50
Gardening, trees early > 2 m	0.04	0.12	0.21	0.32	0.45	0.59	0.74	0.90	1.09	1.29	1.50	1.72
Gardening, trees late	0.00	0.01	0.01	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.07	0.08

Table E7: Proportion of field rate to be covered by a limit test to ensure ≤ 50 % overall effect level for different spray drift scenarios when the field rate = ER_{25}, logit model (figures must be multiplied by the appropriate safety factor)

Number of applications	1	2	3	4	5	6	7	8	9	10	11	12
Arable crops	0.01	0.03	0.05	0.08	0.11	0.14	0.18	0.21	0.25	0.30	0.35	0.40
Arable crops × 2	0.02	0.06	0.11	0.16	0.22	0.28	0.36	0.42	0.50	0.60	0.70	0.80
Arable crops > 900 L/ha	0.01	0.05	0.09	0.10	0.14	0.17	0.21	0.23	0.27	0.32	0.38	0.43
Vines	0.03	0.09	0.19	0.29	0.42	0.55	0.70	0.86	1.04	1.23	1.43	1.65
Orchards, early	0.10	0.33	0.64	1.04	1.47	1.96	2.51	3.06	3.69	4.37	5.09	5.85
Orchards, late	0.05	0.16	0.30	0.44	0.62	0.79	1.01	1.19	1.44	1.70	1.98	2.28
Hops	0.06	0.23	0.43	0.67	0.96	1.28	1.62	1.86	2.25	2.66	3.10	3.56
Knapsack sprayer, plants < 50 cm	0.01	0.03	0.05	0.08	0.11	0.14	0.18	0.21	0.25	0.30	0.35	0.40
Knapsack sprayer, plants > 50 cm	0.03	0.09	0.19	0.29	0.42	0.55	0.70	0.86	1.04	1.23	1.43	1.65
Knapsack sprayer, spraying screen	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02
Gardening, plants < 50 cm	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03
Gardening, plants > 50 cm	0.00	0.01	0.01	0.02	0.02	0.03	0.03	0.04	0.05	0.06	0.06	0.07
Gardening, trees early < 2 m	0.05	0.12	0.20	0.31	0.43	0.54	0.67	0.79	0.96	1.13	1.32	1.51
Gardening, trees early > 2 m	0.13	0.36	0.64	0.97	1.36	1.76	2.21	2.70	3.26	3.86	4.49	5.16
Gardening, trees late	0.01	0.03	0.04	0.06	0.08	0.09	0.11	0.12	0.15	0.18	0.20	0.23

Table E8: Proportion of field rate to be covered by a limit test to ensure ≤ 50 % overall effect level for different spray drift scenarios when the field rate = ER_{50}, logit model (figures must be multiplied by the appropriate safety factor)

Number of applications	1	2	3	4	5	6	7	8	9	10	11	12
Arable crops	0.03	0.09	0.16	0.24	0.33	0.42	0.53	0.63	0.76	0.90	1.04	1.20
Arable crops × 2	0.06	0.19	0.32	0.49	0.67	0.85	1.07	1.25	1.51	1.79	2.09	2.40
Arable crops > 900 L/ha	0.04	0.15	0.28	0.30	0.41	0.51	0.64	0.68	0.82	0.97	1.13	1.29
Vines	0.08	0.28	0.56	0.88	1.26	1.65	2.10	2.58	3.12	3.69	4.30	4.94
Orchards, early	0.29	1.00	1.93	3.11	4.42	5.87	7.53	9.17	11.08	13.12	15.28	17.56
Orchards, late	0.16	0.47	0.89	1.33	1.86	2.38	3.02	3.57	4.31	5.11	5.95	6.84
Hops	0.19	0.69	1.28	2.02	2.89	3.85	4.86	5.58	6.74	7.98	9.30	10.68
Knapsack sprayer, plants < 50 cm	0.03	0.09	0.16	0.24	0.33	0.42	0.53	0.63	0.76	0.90	1.04	1.20
Knapsack sprayer, plants > 50 cm	0.08	0.28	0.56	0.88	1.26	1.65	2.10	2.58	3.12	3.69	4.30	4.94
Knapsack sprayer, spraying screen	0.00	0.00	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.04	0.04	0.05
Gardening, plants < 50 cm	0.00	0.01	0.01	0.02	0.02	0.03	0.04	0.04	0.05	0.06	0.07	0.08
Gardening, plants > 50 cm	0.01	0.02	0.04	0.05	0.07	0.08	0.10	0.12	0.14	0.17	0.19	0.22
Gardening, trees early < 2 m	0.14	0.36	0.60	0.93	1.28	1.62	2.01	2.37	2.87	3.39	3.95	4.54
Gardening, trees early > 2 m	0.38	1.09	1.93	2.91	4.08	5.27	6.64	8.09	9.78	11.58	13.48	15.49
Gardening, trees late	0.04	0.08	0.13	0.18	0.23	0.28	0.32	0.37	0.44	0.53	0.61	0.70

E.5. Discussion of applicability and uncertainties of the effect-based MAF

The effect-based MAF is a theoretically founded concept that is intended to replace the also theoretically founded fate-based MAF in situations in which a carry-over of effects on a target (organism or population) between individual exposure events cannot be excluded and/or where the environmental fate of a test compound is already integrated in the result from a toxicity test.

The approach is built upon two well-established toxicological concepts, namely the probit or logit model for describing dose–response relationships from toxicity experiments and the IA concept of assessing mixture toxicity for describing cumulation of effects following multiple exposure. It is therefore deemed the most appropriate concept for a scenario in which the same individual target organism is (or at least can be) exposed repeatedly, as is the case for NTTPs in the vicinity of agricultural fields.

In contrast, it is the underlying assumption of a fate-based MAF that the effect level after a series of exposure events is identical to the effect level after a single exposure to the remaining exposure level after the last application (taking into account compound degradation). This assumption would, strictly speaking, exclusively correspond to a scenario in which the target organism is exposed only once to the resulting/final concentration level. There would be another possible assumption to justify a fate-based MAF, i.e. that effect levels in NTTPs increase and decrease over time directly proportional to the course of concentration levels of a compound on the plants (determined by its environmental fate parameters). However, there is no scientific justification for such a hypothesis.

To achieve a feasible approach for effect cumulation, simplified assumptions with regard to toxicodynamics had to be made insofar as the effect level observed in the OECD 227 test after the observation period of 21 to 28 days is considered to be reached directly at the time of exposure and to remain constant thereafter. The effect-based MAF would thus address a situation in which the subsequent exposure occurs only at the end of this observation period. True effect levels before and after this point might be higher or lower, depending on actual toxicodynamics in the tested plants, for which no information is available in a standard test. If there is available information indicating that due to an application interval considerably longer than the test duration, effects predicted by the test (e.g. loss of biomass) would have vanished at the time of a subsequent application, the applications can be assessed individually as single applications. For terrestrial plants in off-field areas, it should normally be considered that they are exposed repeatedly to multiple applications of a PPP, because their life span is long enough. Therefore, a justification for assessing multiple applications as individual events would require experimental data to prove that:

a) internal exposure concentrations in the individual organisms will drop below critical threshold levels; and

b) complete repair of damage will occur between the exposure peaks.

As such information is usually not available for terrestrial plants, multiple applications can thus not be assessed individually based on current test methods.

Due to the same lack of information on actual toxicodynamics, it is not possible to make a quantitative or even qualitative prediction whether plants will be more or less susceptible to repeated exposure events compared with the initial exposure event. Against the background that testing of early growth stages in the OECD 227 test is generally considered to constitute a worst case, the assumptions that neither recovery nor increased susceptibility would occur in tested plants over a series of exposure events are considered acceptable for a standard risk assessment method.

However, it must be noted that this simplified consideration of toxicodynamics is not a specific feature of the effect-based MAF concept. The same assumptions are also made in all other standard methods for ecotoxicological risk assessment, i.e. neither the time between exposure and determination of

effect levels in the test nor the differences between test and field conditions and their possible impact on the fate and toxicokinetics of a compound during that period are normally taken into account. A more detailed consideration of toxicokinetics and toxicodynamics in the test compared with the field situation would require additional information and could only be dealt with as a part of a refined risk assessment. Still, as a consequence, applying an effect-based MAF in a risk assessment for NTTPs for application intervals that markedly exceed the observation period of the OECD 227 test is afflicted with increasing uncertainty, since experience suggests that recovery from less severe effects after 21 to 28 days could occur in the tested plants, whereas severe effects might increase even further (but would then already designate an unacceptable situation in regulatory terms).

In summary, it is concluded that, in the case of the NTTP vegetative vigour test, the effect-based MAF makes better use of the available data and is thus deemed scientifically better justified than the fate-based MAF.

Appendix F. Mixture toxicity (How to consider mixtures in the risk assessment)

F.1. Concentration Addition (CA) model

In the following, the CA model is described in a generic way, thus also referring to "concentrations" in general, which is meant to include "rates" (areic concentrations) as well.

F.1.1. Prediction of toxicity values

The CA model is based on the following equation, for deriving a predicted $EC_{x,mix\text{-}CA}$ or $NOEC_{mix\text{-}CA}$ value for a mixture of n (active) substances, present in the mixture at a relative fraction of p_i, with known toxicity ($EC_{x,i}$ or $NOEC_i$), assuming concentration additivity:

$$EC_{x,mix\text{-}CA} = \left(\sum_{n} \frac{p_i}{EC_{x,i}} \right)^{-1}$$

When Σp_i is 1, the result of the calculation represents the expected toxicity of a mixture for which applicable toxicity values are known for each component. In cases in which it can be assumed that some mixture components do not significantly contribute to the toxicity of the mixture (which often also means that no applicable toxicity values are available for those components), then Σp_i could also be smaller than 1. Under such circumstances, the calculated $EC_{x,mix\text{-}CA}$ would thus reflect the toxicity of a mixture composed of toxic and inert compounds, the latter diluting the toxicity of the previous.

F.1.1.1. Toxic unit (TU) concept

According to the European scientific committees (SCHER, SCCS, SCENIHR, 2012), TUs are defined as "the ratio between the concentration (i.e. c_i) of a mixture component and its toxicological acute (e.g. short-term EC_{50}) or chronic (e.g. long-term NOEC) endpoint". In addition, the TU of a mixture has been defined as the sum of the TUs of each individual compound of that mixture:

$$\sum_{n} TU_i = \sum_{n} \frac{c_i}{EC_{x,i}}$$

The committees also noted that the TU approach should refer to specific endpoints and to defined taxonomic groups of organisms but not to the ecosystem as a whole.

The TU approach may be applied for identifying compounds in a mixture that determine overall mixture toxicity to such an extent that neither the mixture as such nor its other components need to be considered in the further assessment. Provided it has been confirmed that the CA model will deliver a reliable estimate of mixture toxicity, a mixture component can be considered the "driver" of overall mixture toxicity when its individual TU_i amounts to $\geq$ 90 % of the sum of TUs of the given mixture.

F.1.1.2. Model deviation ratio for counter-checking calculated against measured mixture toxicity

In order to determine if the toxicity of mixture components was increased (i.e. synergism) or decreased (i.e. antagonism) in the mixture compared with their toxicity as single substances, comparison of a calculated $EC_{x,mix\text{-}CA}$ for the mixture (considering all components for which toxicity data are available) versus a measured $EC_{x,mixture}$ is informative. Such a comparison may also indicate cases in which relevant toxicity contributions of components (e.g. co-formulants in a PPP) occur that were not included in the calculation but had to be considered in a refined calculation (provided that single substance toxicity data are available). The deviation between calculated and measured mixture toxicity is—in line with Belden et al. (2007)—termed the model deviation ratio (MDR):

$$\mathrm{MDR} = \frac{\mathrm{EC}_{x,mix-CA}(\text{modelled value})}{\mathrm{EC}_{x,mixture}(\text{measured value})}$$

In the interpretation of the MDR, the following cases are considered:

- $0.2 \leq \mathrm{MDR} \leq 5$
  Measured and modelled toxicities are considered in agreement. This convention is in line with a proposal currently being made for the authorisation of biocidal products under the auspices of the European Chemicals Agency. In such a case, measured toxicity values ($EC_{x,mixture}$) would take precedence over modelled values ($EC_{x,mix\text{-}CA}$) in a risk assessment.
- $\mathrm{MDR} > 5$
  More than additive (i.e. synergistic) mixture toxicity is indicated. A risk assessment should be based on measured toxicity values ($EC_{x,mixture}$), because synergistic interactions are not predictable either by CA or by other concepts such as IA.
- $\mathrm{MDR} < 2$
  Less than additive (i.e. antagonistic) mixture toxicity is indicated. For precautionary reasons and to cover the possible variability of toxicity test results, a risk assessment should be based on modelled toxicity values ($EC_{x,mix\text{-}CA}$), unless a plausible toxicological explanation for the apparent antagonism can be provided (e.g. special features of a formulation type).

Various obstacles could hamper the interpretation of an MDR, among them a limited availability of adequate toxicity data for individual mixture components, or the heterogeneity of the available toxicity data in terms of tested species, exposure designs, etc. This must be considered in a careful interpretation of an MDR. It must also be ensured that measured and modelled actually refer to the same basis, i.e. the relative proportion of mixture components must be consistent. For liquid mixtures, it might become necessary to also take into account measured density values in estimating the modelled toxicity values.

F.2. Independent Action model

Like the CA model, the IA model is also described in a generic way, referring to "concentrations" in general, which is meant to include also "rates" (areic concentrations).

F.2.1. Prediction of toxicity values

The IA model is based on the two core assumptions: (i) that the toxicity of each of a number of simultaneously acting compounds is not influenced by the presence of the other compounds; and (ii) that all compounds affect the same biological endpoint. According to this concept, the mixture effect level (x_{cum}) for n compounds causing individual effect levels, x_i, is calculated as follows:

$$x_{mix} = 1 - \prod_{n}(1 - x_i)$$

Different from the CA model, the IA model does not directly deliver effect concentrations (EC_x or NOEC values) for use in a standard risk assessment. Instead, an effect level for the mixture is calculated that results from the joint action of the mixture components, each contributing with its own individual effect level.

Linking effect values to concentrations is possible when the dose–response relationship is known for the compounds in the mixture. As an example, the dose–response curve can be described by the probit or the logit model, in which the dependency of the effect level x from the exposure level c is defined by either the distribution function Φ of a normal (0,1) distribution or the logistic function with the

parameters *a* (slope) and *b* (intercept) determining the shape and the central position of the resulting sigmoid curve, respectively.

$$x = \Phi(a \ln c + b) \quad \Leftrightarrow \quad \Phi^{-1}(x) = a \ln c + b \qquad \text{(probit model)}$$

$$x = \frac{1}{1+e^{-(a \ln c + b)}} \quad \Leftrightarrow \quad \ln \frac{x}{1-x} = a \ln c + b \qquad \text{(logit model)}$$

The derivation of equations for calculating the slope and intercept parameters, *a* and *b*, respectively, is described in Appendix E, detailing the consideration of multiple applications in the risk assessment, and is therefore not repeated here in detail. When two effect concentrations, EC_x and EC_y, are known for a compound, the dose–response parameters can be calculated as follows.

$$a = \frac{\Phi^{-1}(x) - \Phi^{-1}(y)}{\ln EC_x - \ln EC_y} \qquad b = \Phi^{-1}(x) - a \ln EC_x \qquad \text{(probit model)}$$

$$a = \frac{\ln \frac{x}{1-x} - \ln \frac{y}{1-y}}{\ln EC_x - \ln EC_y} \qquad b = \ln \frac{x}{1-x} - a \ln EC_x \qquad \text{(logit model)}$$

Recalculation from an effect level, *x*, to the corresponding concentration, EC_x, is possible as follows.

$$\ln EC_x = \frac{\Phi^{-1}(x) - b}{a} \quad \Leftrightarrow \quad EC_x = \exp\left(\frac{\Phi^{-1}(x) - b}{a}\right) \qquad \text{(probit model)}$$

$$\ln EC_x = \frac{\ln \frac{x}{1-x} - b}{a} \quad \Leftrightarrow \quad EC_x = \exp\left(\frac{\ln \frac{x}{1-x} - b}{a}\right) \qquad \text{(logit model)}$$

By calculating the IA model for two different concentrations of a mixture (i.e. considering two concentration-dependent effect levels per mixture component) without changing its relative composition, two mixture effect levels $EC_{x,mix\text{-}IA}$ and $EC_{y,mix\text{-}IA}$ can be obtained, from which, in turn, dose–response data for the mixture can be obtained. Thus, it is also possible to draw conclusions on suitable $EC_{x,mix\text{-}IA}$ values for use in a standard risk assessment.

F.3. Mixture toxicity and risk assessment

F.3.1. General aspects

As already mentioned in the introduction, exposure to mixtures is a frequent and common phenomenon for NTTPs in the vicinity of agricultural fields. This is to some extent already reflected in the data requirements and risk assessment procedures. To address the risk from exposure to PPPs containing one or more active substances in combination with formulants that are capable of and often even intended to have an impact on the toxicity of the product, all toxicity tests must, in principle, be conducted with the formulated product itself. Therefore, in the ideal case, there should be no fundamental lack of data on the toxicity of a PPP to NTTPs. Toxicity data are typically available for six or more species, and further broadening of that database is possible without the ethical limitations that had to be considered for animal testing.

Nevertheless, even with good available data on product toxicity, questions could remain for the risk assessment that would require additional input from calculation or modelling tools. The most obvious example are tank mixtures for which specific authorisations are sought. While it appears logical to demand the same type and number of data for such tank mixes as for formulated products with several active substances, no such requirements are legally fixed at the moment. Therefore, tools for modelling mixture toxicity might be applied to underpin, where appropriate, an argument for demanding data on a case-by-case basis or to provide appropriate toxicity estimates for a risk assessment.

A second possible reason for applying tools for modelling mixture toxicity is linked to the high diversity of the plant realm. A broad spectrum of plant species can be used for testing and their sensitivity to certain active substances or PPPs can differ markedly. It may thus happen in the evaluation of a certain product that datasets for other formulations with the same active substance or even the same combination of active substances as the product under assessment could include test results for sensitive species that were not considered in the spectrum of test species for the evaluated product. Again, those tools might be applied for the analysis of such data, to decide whether the database of the evaluated product could or should be broadened, and they might also be used for the prediction of toxicity values where appropriate.

F.3.2. Starting points and important aspects for considering mixture toxicity

As mentioned in the introduction, there are different scenarios with regard to the evaluation of mixture effects in the risk assessment for NTTPs. Mixture toxicity models can be applied as analytical tools to support the interpretation of experimental data, but they may also be used for the prediction of toxicity levels in certain cases. In the following section, a short overview is provided on possible starting points and important aspects for considering mixture toxicity.

F.3.2.1. Assessment of products containing one active substance

NTTP toxicity tests with the product are required and constitute the basis for risk assessment.

Are toxicity data available for other products containing the same active substance (EU active substance evaluation, data from authorisation procedures)?
Do other data indicate a significantly higher toxicity to plant species that were not tested with the product to be assessed?

⇒ If yes, check potential for extrapolation.

Data from efficacy testing should always be analysed.

Are there indications that sensitive groups, as identified in the efficacy assessment, are not sufficiently represented in the ecotoxicology dataset (consider information on mode of action and experimental data)?

⇒ If yes, check potential for extrapolation.

F.3.2.2. Assessment of products containing more than one active substance

NTTP toxicity tests with the product are required and constitute the basis for risk assessment.

Are toxicity data available for other products containing one or several of the active substances (EU active substance evaluation, data from authorisation procedures)?
Do other data indicate a significantly higher toxicity to plant species that were not tested with the product to be assessed?

⇒ If yes, check potential for extrapolation and/or toxicity modelling.

Data from efficacy testing should always be analysed

Are there indications that sensitive groups as identified in the efficacy assessment are not sufficiently represented in the ecotoxicology dataset (consider information on mode of action and experimental data)?

⇒ If yes, check potential for extrapolation and/or toxicity modelling.

F.3.2.3. Assessment of tank mixtures

NTTP toxicity tests with the tank mixture are not explicitly required but would clearly constitute the most reliable basis for risk assessment.

Are toxicity tests with the tank mixture available in sufficient number and of sufficient quality?

⇒ If yes, proceed as for the assessment of products containing more than one active substance.

⇒ If no, check the quality of data for all products in the mixture, as described above, and then check potential for toxicity modelling.

Data from efficacy testing should always be analysed.

Are there indications that sensitive groups, as identified in the efficacy assessment, are not sufficiently represented in the ecotoxicology dataset for the tank mixture, if available, or the products in the mixture (consider information on mode of action and experimental data)?

⇒ If yes, check potential for toxicity modelling.

F.3.2.4. Checking the potential for extrapolation

Extrapolation of toxicity information from an active substance or from another product to the product to be assessed is possible only when the impact of other formulants is either negligible or can be quantified with sufficient certainty. The MDR approach can be used for estimating the impact of formulants that are not toxic themselves but would have a significant impact ton he toxicity of active compounds. This is also possible for “mixtures” containing only one active compound. All available data, including information from efficacy testing and other reliable sources, should be included in the considerations.

Is the impact of other formulants on the toxicity of the compared products either negligible or capable of being quantified with sufficient certainty?

⇒ If yes, extrapolation of toxicity information is possible. (Still, it should be noted that additional experimental data will normally take precedence over calculation of results in the risk assessment.)

⇒ If no, further testing of the product under assessment on sensitive species is necessary.

F.3.2.5. Checking the potential for toxicity modelling

With regard to formulants that are not toxic themselves, the same criteria as for extrapolation of toxicity information also apply for toxicity modelling. Beyond that, toxicity modelling can provide reliable results only when no synergistic interaction of active compounds occurs. Again, the MDR approach can be used for estimating the impact of formulants that are not toxic themselves. With respect to the potential for synergism of active compounds, not only the results of NTTP tests from the ecotoxicology dataset but also the information from efficacy testing, in particular on the mode of action and on more than additive effects on target plants, should be carefully evaluated. In addition, data from other reliable sources should be included in the considerations.

Is the impact of other formulants on the toxicity of the compared products either negligible or capable of being quantified with sufficient certainty?

⇒ If yes, toxicity modelling may be possible—see below. (Still, it should be noted that additional experimental data will normally take precedence over calculation of results in the risk assessment.)

⇒ If no, further testing of the product under assessment on sensitive species is necessary.

If the impact of other formulants on the toxicity of the compared products is either negligible or capable of being quantified with sufficient certainty, what kind of effect and effect combination is observed or can be reliably deduced from the available information?

⇒ If more than additive effects are observed or their occurrence is likely, owing to the available mechanistic data or other experimental evidence, further testing of the product under assessment on sensitive species is necessary.

⇒ If cumulation of effects is observed or its occurrence is likely, owing to the available mechanistic data or other experimental evidence, or if the type of interaction between active compounds cannot be clarified, the CA model should be used for toxicity modelling. (Still, it should be noted that additional experimental data will normally take precedence over calculation of results in the risk assessment.)

⇒ If there is enough evidence that the core assumptions of the IA model (toxicity of each compound is not affected by the presence of other compounds and all compounds affect the same biological endpoint) is valid for the product to be assessed, the IA model may be used for toxicity modelling. This could, for example, be the case where the different active compounds obviously affect different groups of plants. (Still, it should be noted that additional experimental data will normally take precedence over calculation of results in the risk assessment.)

F.3.3. Specific issues for NTTPs

F.3.3.1. Risk assessment based on measured data

The risk for NTTPs in the vicinity of agricultural fields is assessed by calculating TER values as the quotient of an appropriate toxicity estimate (ER_x) and a predicted environmental rate due to deposition of spray drift (and, where relevant and taken into account in the assessment, volatilisation) from the treated field. Where the assessment can be focused on spray drift entries from a single product, no specific consideration of mixture toxicity is necessary at this stage.

If a product contains a (semi-)volatile besides non-volatile active substances, assessors might wish to consider a changed ratio of active substances in the off-field deposits compared with the product. Provided that toxicity data are available for all active substances, the predicted toxicity (applying the CA model by default or the IA model where scientifically justified) of the mixture of active substances in the product can be compared with predicted toxicity of the mixture in the off-field deposits. If the deviation between those two calculated values is less than 20 %, the risk assessment can be performed with the toxicity data for the product. If this deviation exceeds 20 %, the changed ratio of active substances in the off-field deposits would have to be considered case by case in a refined risk assessment.

F.3.3.2. Consideration of toxicity values for different test species

Different from other areas of the environmental risk assessment, the standard assessment for NTTPs is based on toxicity data from several test species already at the screening stage. When it comes to quantitative evaluation, either aggregated toxicity values (e.g. the HC_5 from an SSD as the preferred option) or worst-case values from the available multi-species dataset are used for the TER calculation. Strictly speaking, this is not in agreement with the basic assumptions of the mixture toxicity models that were originally derived for the prediction of toxicity estimates for one defined target species. However, for the purpose of environmental risk assessment in general and for the NTTP assessment in particular, it should be taken into account that the species used in standard testing do not constitute representatives for the protection goal on species level, but, in a more abstract way, in terms of possible sensitivity. It is therefore deemed appropriate to consider an HC_5 (calculated from a set of ER_x values) or an ER_x of the most sensitive tested species as a proxy for the ER_x of "plants" and use them accordingly in mixture toxicity calculations.

F.3.3.3. Simplified application of the CA model on TER values

Taking into account the conclusion that an HC_5 or a lowest ER_x obtained for a PPP can be considered a proxy for the ER_x of "plants" for this product, a simplified approach for the quantitative risk assessment can be proposed. TER values for a mixture are calculated as the quotient of the $ER_{x,mix}$ and the predicted environmental (deposition) rate of the whole mixture (PER_{mix}). When concentration additivity of toxicity is assumed, the $ER_{x,mix}$ can be calculated with the CA model, while the PER_{mix} can be expressed as the sum of individual PER_i values, each defined as $p_i \times PER_{mix}$ (with the same fractions p_i as in the CA equation).

$$TER_{mix} = \frac{ER_{x,mix}}{PER_{mix}} = \frac{\left(\sum_n \frac{p_i}{ER_{x,i}}\right)^{-1}}{\sum_n PER_i} = \left(\sum_n \frac{p_i \times \sum_n PER_i}{ER_{x,i}}\right)^{-1}$$

In the standard case, the NTTP risk assessment makes use of initial drift deposits, in which the relative composition of active substances is thus identical to the composition of the mixture applied to the field (single product or tank mixture). Thus, the fractions p_i can also be substituted by $p_i = PEC_i / \Sigma PEC_i$ in the equation, which then shows that the TER_{mix} for the mixture can be calculated directly from the TER_i values for the individual mixture components.

$$TER_{mix} = \left(\sum_{n} \frac{PER_i}{ER_{x,i}} \right)^{-1} = \left(\sum_{n} \frac{1}{TER_i} \right)^{-1}$$

This method of calculation is in particular recommended for the assessment of tank mixtures, as long as no synergism is expected. Where the TER acceptability criterion has been adjusted for one of the mixture components in the context of a refined assessment, the actual TER values for the mixture components must be divided by the respective acceptability criteria before the mixture calculation is made. In other words, the mixture toxicity assessment will then be performed on the basis of regulatory acceptable concentrations, i.e. those concentrations that may not be exceeded under field conditions.

GLOSSARY AND ABBREVIATIONS

AF	Assessment Factor
a.s.	Active substance
CA	Concentration addition
DG SANCO	Directorate General for Health and Consumers
EC_x	Concentration at which *x* % effect was observed/calculated
Efficacy data and crops' margin of safety data	Data on the efficacy of the a.s. on the pest; data on adjacent crops, data on succeeding crops
EFSA	European Food Safety Authority
ER	Effect Rate
EU	European Union
FOCUS	FOrum for the Co-ordination of pesticide fate models and their USe
GD	Guidance Document
HC_x	Hazardous concentration for *x* % of the species of an SSD
IA	Independent Action
$LLHC_5$	Lower limit of the confidence interval of the hazardous concentration for 5 % of the species of an SSD
LOEC	Lowest observed effect concentration
MAF	Multiple Application Factor
NOEC	No observed effect concentration
NOER	No observed effect rate
NTTP	Non-target terrestrial plant
Metabolite	Any metabolite or degradation product of an active substance, safener or synergist, formed either in organisms or in the environment (thus also including oxidation products that may have a larger molecular mass than the parent substance)
OECD	Organisation for Economic Co-operation and Development
PEC	Predicted environmental concentration

PPP	Plant protection product
PPR Panel	EFSA's Panel on Plant Protection Products and their Residues
Pre-screening data	Data collected by the applicant at a very early stage of the development of the compound, i.e. the data on which the company decides whether the compound is worthwhile developing for use in agriculture
RA	Risk assessment
RAC	Regulatory acceptable concentration
SCFCAH	Standing Committee on the Food Chain and Animal Health
Screening data for herbicidal activity	Sometimes called also tier I screening data; screening-data for non-herbicides/active substances without herbicidal or plant growth regulatory activity, as indicated in the data requirements (Commission Regulations (EU) No 283/2013 and No 284/2013)
Semelparous	A semelparous species is characterised by a single reproductive episode during its life
SETAC	Society for Environmental Toxicology and Chemistry
SPG	Specific protection goal
SSD	Species sensitivity distribution