STATEMENT

efsa JOURNAL

ADOPTED: 27 March 2023

doi: 10.2903/j.efsa.2023.7990

Statement of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on the design and conduct of groundwater monitoring studies supporting groundwater exposure assessments of pesticides

EFSA Panel on Plant Protection Products and their Residues (PPR),
Antonio Hernandez-Jerez, Paulien Adriaanse, Annette Aldrich, Philippe Berny, Tamara Coja, Sabine Duquesne, Andreas Focks, Marina Marinovich, Maurice Millet, Olavi Pelkonen, Silvia Pieper, Christopher Topping, Anneli Widenfalk, Martin Wilks, Gerrit Wolterink, Roy Kasteel, Konstantin Kuppe and Aaldrik Tiktak

Abstract

Groundwater monitoring is the highest tier in the leaching assessment of plant protection products in the EU. The European Commission requested EFSA for a review by the PPR Panel of the scientific paper of Gimsing et al. (2019) on the design and conduct of groundwater monitoring studies. The Panel concludes that this paper provides many recommendations; however, specific guidance on how to design, conduct and evaluate groundwater monitoring studies for regulatory purposes is missing. The Panel notes that there is no agreed specific protection goal (SPG) at EU level. Also, the SPG has not yet been operationalised in an agreed exposure assessment goal (ExAG). The ExAG describes which groundwater needs to be protected, where and when. Because the design and interpretation of monitoring studies depends on the ExAG, development of harmonised guidance is not yet possible. The development of an agreed ExAG must therefore be given priority. A central question in the design and interpretation of groundwater monitoring studies is that of groundwater vulnerability. Applicants must demonstrate that the selected monitoring sites represent realistic worst-case conditions as specified in the ExAG. Guidance and models are needed to support this step. A prerequisite for the regulatory use of monitoring data is the availability of complete data on the use history of the products containing the respective active substances. Applicants must further demonstrate that monitoring wells are hydrologically connected to the fields where the active substance has been applied. Modelling in combination with (pseudo)tracer experiments would be the preferred option. The Panel concludes that well-conducted monitoring studies provide more realistic exposure assessments and can therefore overrule results from lower tier studies. Groundwater monitoring studies involve a high workload for both regulators and applicants. Standardised procedures and monitoring networks could help to reduce this workload.

Keywords: groundwater, pesticides, monitoring, vulnerability assessment, FOCUS

Requestor: European Commission

Question numbers: EFSA-Q-2021-0788, EFSA-Q-2022-00034

Correspondence: pesticides.peerreview@efsa.europa.eu

Panel members: Antonio Hernandez-Jerez, Paulien Adriaanse, Annette Aldrich, Philippe Berny, Tamara Coja, Sabine Duquesne, Andreas Focks, Marina Marinovich, Maurice Millet, Olavi Pelkonen, Silvia Pieper, Aaldrik Tiktak, Christopher Topping, Anneli Widenfalk, Martin Wilks and Gerrit Wolterink.

Declarations of interest: If you wish to access the declaration of interests of any expert contributing to an EFSA scientific assessment, please contact interestmanagement@efsa.europa.eu.

Acknowledgements: The Scientific Panel on Plant Protection Products and their Residues (PPR Panel) wishes to thank the following for the support provided to this scientific output: Laura Padovani, Chris Lythgo, Laia Herrero Nogareda and Mark Egsmose [staff members or others who contributed but are not eligible as authors].

Suggested citation: EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Hernandez-Jerez A, Adriaanse P, Aldrich A, Berny P, Coja T, Duquesne S, Focks A, Marinovich M, Millet M, Pelkonen O, Pieper S, Topping C, Widenfalk A, Wilks M, Wolterink G, Kasteel R, Kuppe K and Tiktak A, 2023. Statement of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on the design and conduct of groundwater monitoring studies supporting groundwater exposure assessments of pesticides. EFSA Journal 2023;21(5):7990, 108 pp. https://doi.org/10.2903/j.efsa.2023.7990

ISSN: 1831-4732

EUROPEAN FOOD SAFETY AUTHORITY

The EFSA Journal is a publication of the European Food Safety Authority, a European agency funded by the European Union.

Summary

The protection of groundwater and the maintenance of its quality is a cornerstone of EU-legislation and therefore a leaching assessment is required for the approval of active substances and the authorisation of plant protection products (PPPs) as laid down in Regulation (EC) No 1107/2009. Groundwater monitoring is a higher-tier option in the leaching assessment of PPPs in the EU. The European Commission requested the European Food Safety Authority (EFSA) for a review of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on the scientific paper by Gimsing et al. (2019) regarding the design and conduct of groundwater monitoring studies supporting pesticide groundwater exposure assessments. Gimsing et al. (2019) provide many recommendations on how to design and conduct groundwater monitoring studies without giving specific guidance on how to design, conduct and evaluate groundwater monitoring studies for regulatory purposes. However, the information provided in Gimsing et al. (2019) can serve as a basis for developing detailed and harmonised guidance for groundwater monitoring studies.

Specific protection goals

Before detailed and harmonised guidance can be developed, risk managers must define which ecosystem services need to be protected where and when (this is called the specific protection goal or SPG). This definition is crucial for designing appropriate risk assessment schemes in which ecotoxicological effects can be combined with exposure measured or simulated in the field. The current leaching assessment scheme for groundwater aims primarily at the ecosystem service 'provisioning of clean drinking water' and does not consider the impacts on non-target organisms. However, when defining SPGs for groundwater, further relevant ecosystem services must be considered, especially those addressing groundwater as habitat for non-target species and the impact of groundwater contamination on surface water and its communities. The Panel recommends clearly defining all SPGs, including all relevant ecosystem services. If the SPGs consider groundwater organisms, an effect assessment scheme should be developed as well. An additional question here is whether the risk assessment for non-target organisms covers biodiversity in groundwater sufficiently well, via e.g. testing of sensitive groups of aquatic organisms and/or soil organisms.

The exposure assessment goal

The exposure assessment is an integral part of any risk assessment scheme. To define which exposure should be used to evaluate the SPG for groundwater in a structured and unambiguous way, the so-called exposure assessment goal (ExAG) must be defined. This allows answering questions such as (i) where, in which environmental compartment, and for which time frame should exposure be estimated, or (ii) how conservative should the exposure estimate be, i.e. what percentage of the exposure situations in the field should be covered in the risk assessment? The Panel notes that a clearly defined ExAG for groundwater leaching is missing at EU level. Because the design of monitoring studies depends on the ExAG, the development of harmonised guidance for the design and conduct of groundwater monitoring studies is not yet possible. The definition of the ExAG should therefore be given priority.

The Panel recommends that the definition of the ExAG is to be done by risk assessors with input from risk managers. Input is particularly needed on the following three elements of the ExAG, because these determine the strictness of the leaching assessment and thus go beyond the mandate of risk assessors:

- The depth relative to the soil surface and/or relative to the saturated zone. Monitoring depth is a compromise between practicability and conservativeness. The concentration in the groundwater generally decreases with depth due to processes such as degradation during downward transport, or dilution with uncontaminated water from non-treated areas. The Panel suggests monitoring in the uppermost groundwater (1–2 m filter screen installed near the groundwater table but shallower than the depth defined by the ExAG), because so obtained groundwater concentrations represent in a conservative way the concentration in the entire groundwater body or aquifer.
- The time window over which concentrations should be averaged. This time window determines whether temporary exceedances of the target value may be permitted. Options range from annual arithmetic mean values to instantaneous maximum values.

- The desired overall level of protection by selecting a percentage of situations that should be covered by the leaching assessment. The Panel notes that the selection of this percentile is to be done by risk managers, because it involves political and socio-economic considerations that are outside the remit of the Panel.

Consistency of the tiered approach

Groundwater monitoring is the highest tier (i.e. Tier-4) of the FOCUS groundwater assessment scheme. According to the principles of a tiered approach, higher tiers should provide a less conservative and a more realistic estimate of the target quantity. The Panel notes that well performed monitoring studies provide a more realistic exposure assessment than the lower (modelling) tiers. If the monitoring depth is greater than 1 m (which is the evaluation depth for the models used at the lower tiers), measured concentrations will generally also be lower than the concentrations obtained in the lower tiers. This can, however, not *always* be guaranteed because processes that are not included in the models can result in higher groundwater concentrations. Because well-conducted monitoring studies provide more realistic exposure assessments, the Panel concludes that a (well-conducted) Tier-4 study can overrule results from lower tier studies.

Preferential flow is one of the known processes that could make the assessment at Tier-4 more conservative. The Panel recommends investigating if it is possible to include a harmonised description of preferential flow in the leaching models that are currently used in the leaching assessment and to investigate the consequences for the groundwater concentration in the aquifer. Ultimately, preferential flow could be incorporated into the lower tiers of FOCUS groundwater.

Vulnerability assessment

The Panel agrees with Gimsing et al. (2019) that groundwater vulnerability is a central question in the design and interpretation of groundwater monitoring studies. Applicants must demonstrate that the selected monitoring sites represent realistic worst-case conditions as specified in the ExAG. For this purpose, it is necessary to model groundwater vulnerability in the entire area of use of the pesticide. It was demonstrated in several studies that substance properties play an important role in the vulnerability of groundwater to pesticide leaching. The consequence is that leaching models are needed to map groundwater vulnerability. Both process-based simulation models and statistical metamodels can be used for this purpose; the Panel considers so-called index approaches inappropriate. The Panel recommends developing harmonised models for vulnerability mapping. This should include the development of harmonised geodata. The models and data sets should be well-documented, brought under version control and preferably be accessible through a user-friendly interface. The geodata should be regularly updated to accommodate for the effects of climate change and land-use change.

The Panel recommends using the annual mass flux at 1-m depth as an indicator of groundwater vulnerability. Given the lack of data on deeper soil layers, and because groundwater vulnerability is to a large extent determined by processes occurring in the topsoil, the Panel considers this the only practical way forward. The Panel notes, however, that vulnerability of the groundwater aquifer may differ from the groundwater vulnerability at 1-m depth (i.e. leaching vulnerability). This gives uncertainty in the vulnerability assessment, which must be considered when selecting the monitoring sites and interpreting results at a later stage.

The Panel notes that the groundwater vulnerability assessment serves two purposes, i.e. (i) identification of potential monitoring sites and (ii) setting monitoring studies into a larger spatial context. The examples given in Gimsing et al. (2019) are illustrative; however, the Panel notes that binding guidance for regulatory use needs to be developed. The Panel agrees with Gimsing et al. (2019) that there is currently no available data set that can be used to identify monitoring sites at *field-level* across the EU. So, the Panel proposes to clearly distinguish between these two phases of site selection, i.e. selection of vulnerable areas and/or regions, and selection of the actual monitoring sites within these vulnerable areas. In the second phase, applicants must demonstrate that the actual site characteristics (e.g. soil and climate data) are comparable to those derived from the vulnerability assessment. Criteria need to be set when these characteristics are sufficiently equal.

Although groundwater vulnerability is substance dependent, the Panel accepts to base the selection of monitoring sites on the substance with the most important concentration exceedances. Whether the selected sites are sufficiently conservative for other substances must be demonstrated in the context setting procedure thereafter. The in-context setting procedure can be used to demonstrate a safe use in different FOCUS zones than the FOCUS zone(s) where the monitoring has been performed. The in-context procedure can also be used for the authorisation process of PPPs at the national level, if requirements set by national authorities are met.

Study design of targeted monitoring programmes

Gimsing et al. (2019) describe two types of targeted monitoring programmes, i.e. in-field and edge-of-field study designs. In-field study designs generally profit from the ease of proving hydrological connectivity between the monitoring wells and the treated fields, which is essential when monitoring is used for regulatory purposes. The disadvantage of in-field study designs is the risk of groundwater contamination due to the installation of the wells, rather than leaching. In this respect, edge-of-field study designs are the preferred option. To safeguard collection of maximum residues in the groundwater originating from the adjacent field, the Panel recommends collecting samples in the uppermost shallow groundwater (1- to 2-m filter screen installed near the groundwater table). This would also be in line with the proposed ExAG.

A prerequisite for the regulatory use of monitoring data is the availability of complete data on the use of the products containing the respective active substances. The application data must be reported in sufficient detail for the whole monitoring period and the relevant preceding time span (considering the travel times of the substance). Scaling is needed if the actual application rate differs from the definition of uses according to Good Agricultural Practices (GAPs). Guidance must be developed on how to perform this scaling. The Panel further recommends developing guidance how to relate the actual application frequencies observed in monitoring studies to the definition of the uses according to the GAP.

As mentioned above, it is crucial to demonstrate hydrological connectivity between the field(s) that received the application (and for which data on use is documented) and the sampling well. Connectivity can be proved using measurements of tracers (e.g. bromide) or pseudo-tracers (e.g. metabolites). An alternative is to simulate pesticide transport from the treated field to the well with a combination of a leaching model and a groundwater model (this is referred to as 'time-of-flight modelling'). The Panel considers the combined use of measurements of pseudo-tracers (e.g. metabolites) or tracers (e.g. bromide) with site-specific time-of-flight modelling exercises to be a suitable and practical approach to demonstrate connectivity and expected travel times. The Panel considers time-of-flight modelling a compulsory part of the analysis, because pseudo-tracers and tracers may have shorter travel times than the substance of concern. The Panel notes that time-of-flight modelling requires a high degree of experience with hydrological modelling and site-specific model parameterisation. Further guidance and training courses would be needed if these approaches become more important in regulatory use.

The Panel recommends developing a statistical approach to determine the minimum number of sites required for regulatory groundwater monitoring. The Panel notes that site-selection involves a high workload for both regulators and applicants. A network of standardised monitoring sites could therefore help regulatory acceptance of groundwater monitoring studies. The Panel recommends investigating if such a monitoring network could be established at EU level.

Publicly available monitoring programmes

Monitoring data collected by third party organisations, such as environmental agencies, water companies and research institutions, are designed for purposes other than for the authorisation of PPPs under Regulation (EC) 1107/2009. They provide important information on the current state of and possible trends in groundwater bodies and provide additional information on the applicability and interpretation of targeted monitoring studies due to vast amount of information that is generally available. Therefore, the Panel recommends including publicly available monitoring data in the risk assessment as supportive information. If the monitoring data shows, for example, large-scale exceedances of the drinking water limit for a substance that has passed the lower tiers of the groundwater leaching assessment, risk mitigation measures might be implemented and applicants could be asked to provide additional information to demonstrate the safe use of the product.

At present, publicly available monitoring data often do not meet the main quality criteria for a Tier-4 risk assessment. The Panel recommends using these data as supportive information in a weight-of-evidence approach and not to use these data to overrule earlier tiers. To include publicly available monitoring data as supportive information in the risk assessment, the data need to comply with quality criteria. These quality criteria must be established in a future guidance document. Guidance is further needed on the interpretation of monitoring data of shared metabolites. Access to the data for regulators could be improved by making them available in a central database, together with the metadata and procedures to establish a plausible relationship between the presence of the substance in groundwater and the authorised use.

Reporting results from groundwater monitoring programmes

Gimsing et al. (2019) give some considerations for the study report of monitoring studies to be submitted to regulatory authorities. A list of items to be reported is included as well. This list is, however, not comprehensive for targeted monitoring studies. The Panel suggests creating a fixed list of items to be reported. The monitoring report must contain at least the following elements:

- A description of the set-up of the monitoring study, including results from the site selection and the in-context setting procedure,
- Monitoring results from all individual monitoring sites including a critical evaluation of results,
- A description of the aggregation procedure to derive the regulatory endpoint, including an evaluation of results,
- A discussion section, addressing the most important uncertainties of the study,
- A conclusion. The study report should include whether a safe use can be identified for the intended use or not. This conclusion should be based on an evaluation of the available data against the agreed ExAG.

The discussion section must contain an analysis of the most important uncertainties from the monitoring study. A balanced discussion of false positives and false negatives must be part of this analysis. The Panel suggests that guidance needs to be developed on how to use results from this uncertainty analysis in regulatory assessments.

Table of contents

Abstract 1
Summary 3
1. Introduction 8
1.1. Terms of reference as provided by the requestor 8
1.2. Interpretation of the terms of reference by the Panel 9
2. Specific protection goals and exposure assessment goals for leaching to groundwater 9
2.1. Groundwater in support of and as a habitat for biodiversity 9
2.2. Need for defining both specific protection goals, and exposure assessment goals 11
2.4. Exposure assessment goal for groundwater 13
2.5. Review of the exposure assessment options in Gimsing et al. (2019) 17
2.6. Conclusions and recommendations 19
3. Principle of the tiered approach and challenges for the groundwater assessment 19
3.1. The type of concentration 20
3.2. Processes that may affect the conservativeness of Tier-4 22
3.3. Conclusions and recommendations 24
4. Vulnerability assessment for site selection and context setting 24
4.1. Consistency with the proposed exposure assessment goals 24
4.2. Intrinsic and specific vulnerability 26
4.3. Suitable models for vulnerability mapping 26
4.4. Use of vulnerability mapping for site selection and context setting 28
4.4.1. Identification of potential monitoring sites 28
4.4.2. In-context setting of monitoring sites 29
4.5. Conclusions and recommendations 31
5. Data quality considerations 32
5.1. General issues 32
5.2. Conclusions and recommendations 33
6. Study design of targeted monitoring programmes 33
6.1. Representative study designs 34
6.2. History of use 34
6.3. Connectivity 36
6.4. Sampling frequency 40
6.5. Number of monitoring sites 40
6.6. Fixed monitoring networks 41
6.7. Conclusions and recommendations 42
7. Requirements for using public monitoring programmes 43
7.1. Quality criteria 43
7.2. Shared transformation products 43
7.3. Conclusions and recommendations 44
8. Reporting results from groundwater monitoring studies 44
8.1. Description of the set-up of the monitoring study 45
8.2. Reporting of results from the individual monitoring sites 46
8.3. Derivation of regulatory endpoints 47
8.4. Discussion and conclusions from the monitoring study 47
8.5. Conclusions and recommendations 48
9. General conclusions and recommendations 48
9.1. The exposure assessment goal 49
9.2. The tiered approach 49
9.3. The vulnerability assessment 50
9.4. Study design of targeted monitoring programmes 50
9.5. Publicly available monitoring programmes 51
9.6. Reporting results from groundwater monitoring programmes 51
References 52
Abbreviations 56
Appendix A – Resident vs flux concentrations 58
Appendix B – Indicators for vulnerability assessment 62
Appendix C – List of items to characterise publicly available monitoring data 68
Appendix D – Outcome of the public consultation 72

1. Introduction

The European Commission requested the European Food Safety Authority (EFSA) for a Statement of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on the scientific paper by Gimsing et al. (2019) regarding the design and conduct of groundwater monitoring studies supporting pesticide groundwater exposure assessments. In preparation of this Statement, a public consultation was organised by EFSA, and comments were considered when preparing this Statement. The Statement first reviews the options for the Exposure Assessment Goal (ExAG) described in Gimsing et al. (2019). The Panel considers this a crucial starting point, because without a clearly defined ExAG it is not possible to give appropriate guidance. The following sections review the position of groundwater monitoring in the tiered approach and the assessment of groundwater vulnerability supporting the selection and interpretation of groundwater monitoring studies. Then, scientific and technical aspects of study designs for both targeted and public monitoring programmes are discussed. The final section gives some considerations for the monitoring report. Results of the public consultation including the comments received and answers to each comment are published in Appendix D.

1.1. Terms of reference as provided by the requestor

The European Commission requested EFSA for a Statement of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on the design and conduct of groundwater monitoring studies supporting pesticide groundwater exposure assessments. Following Terms of Reference were provided by the requestor.

The protection of groundwater and the maintenance of its quality is a cornerstone of EU legislation and is reflected in the criteria for the approval of active substances and the authorisation of plant protection products (PPPs) as laid down in Regulation (EC) No 1107/2009[1,2]. The data requirements set in point 7.5 of the Commission Implementing Regulation (EU) No 283/2013[3] provide that available monitoring data concerning active substance and relevant metabolites, breakdown and reaction products in soil, groundwater, surface water, sediment and air shall be reported.

The difference among ad-hoc monitoring (targeted) and surveillance is stated in a guidance under the Sustainable Use Directive (EC, 2017). Other available guidance in the FOCUS groundwater report (EC, 2014) addresses performing monitoring studies regarding groundwater and how to consider both targeted monitoring and publicly available monitoring data in the groundwater exposure assessment. Developments on this topic have occurred since 2014.

During the meeting of the Pesticides Steering Network (PSN) held on 17–19 March 2020,[4] Austria and other Member States (France, Germany, the Netherlands and Sweden) suggested to move towards a more harmonised use of groundwater monitoring data in EU regulatory pesticide assessments. The PSN noted that new scientific recommendations were developed (Gimsing et al., 2019) on how to design and conduct groundwater monitoring studies.

Member States were also consulted via the Standing Committee on Plants, Animals, Food and Feed (SCoPAFF) during the phytopharmaceuticals-legislation meeting held in May 2020.[5] Member States called on the Commission to mandate EFSA to organise a public consultation on the publication of Gimsing et al. (2019) and to deliver a statement on this publication considering also the inputs coming from the public consultation. The scientific and technical aspects of study designs and procedures as well as criteria on the assessment of the reliability/quality of the groundwater monitoring information should be addressed, in view of using this information in the EU regulatory exposure assessment of pesticides in line with the requirements set in Commission Implementing Regulation (EU) No 283/2013. Ultimately, guidance for applicants and risk assessors on best practice in this regard should be given.

---

[1] Regulation (EC) No 1107/2009 concerning the placing of plant protection products on the market (OJ L 309, 24.11.2009, pp. 1–50). http://data.europa.eu/eli/reg/2009/1107/oj.

[2] In the European Union (EU) the current regulatory scheme for the assessment of the leaching potential of active substances and their metabolites to groundwater following the use of PPPs is based on a tiered approach, with each tier having the same objective of ensuring the protection of groundwater and ensuring that the objectives laid down in the Water Framework Directive, the Groundwater Directive and the Drinking Water Directive are fulfilled:Directive 2006/118/EC, OJ L 372, 27.12.2006, pp. 19–31 and Directive (EU) 2020/2184, OJ L 435, 23.12.2020, pp. 1–62.

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

[4] https://www.efsa.europa.eu/sites/default/files/event/2020/26th-meeting-efsa-pesticide-steering-network-minutes.pdf

[5] https://ec.europa.eu/food/sites/food/files/plant/docs/sc_phyto_20200518_ppl_sum.pdf

Considering the above, the Commission requested EFSA to arrange a public consultation on Gimsing et al. (2019) followed by a review by the PPR Panel considering also the comments received via the consultation and to provide a statement under Article 29 of Regulation (EC) No 178/2002 on these matters.

1.2. Interpretation of the terms of reference by the Panel

The European Commission requested a review of the scientific paper by Gimsing et al. (2019), focusing mainly on scientific and technical aspects of study designs. EFSA performed a public consultation of the scientific paper by Gimsing et al. (2019) on the design and conduct of groundwater monitoring studies supporting pesticide groundwater exposure assessment from 26 January to 8 March 2022. The Panel[6] interpreted the European Commission-request broadly by also addressing aspects of study design that are missing in the paper but are considered important for developing a guidance document on the conduct and interpretation of groundwater monitoring studies for regulatory purposes. Although evaluation of Specific Protection Goals (SPGs) and/or ExAG were not explicitly mentioned in the Terms of Reference, the Panel reviewed the options for the ExAG that operationalise the SPGs (e.g. the temporal dimension of the relevant type of concentration to assess) and consistency within the FOCUS tiered approach. The Panel – supported by many comments from stakeholders – considered this review necessary, so that risk managers can make an informed decision on SPGs and ExAGs. The availability of clearly defined SPGs and ExAGs is a crucial starting point for any guidance development. The Panel notes that the final selection of these goals is to be done by risk managers, because it involves political and socio-economic considerations that are outside the remit of the Panel. The Terms of Reference do not request an evaluation of field leaching studies. The Panel notes, however, that the distinction between field leaching studies and groundwater monitoring studies in Gimsing et al. (2019) is not always clear and therefore field-leaching studies are discussed in Section 6. Finally, the European Commission requested that ultimately guidance for applicants and risk assessors should be given. It must be noted that this scientific statement is not a guidance document; it is prepared as a *precursor* towards such a guidance document.

2. Specific protection goals and exposure assessment goals for leaching to groundwater

As indicated in Gimsing et al. (2019), clearly defined SPGs for groundwater are missing at EU level. The Panel considers clearly defined SPGs a crucial starting point for the development of a guidance document. In this section, the Panel gives considerations for risk managers when selecting SPGs for groundwater. Following Gimsing et al. (2019), the Panel defines groundwater as 'all water which is below the surface of the ground in the saturation zone and in direct contact with the ground or subsoil'. This implies that temporary zones of perched water are not included.

This section addresses the importance of groundwater biodiversity for several ecosystem services (Section 2.1). Then, the framework for establishing SPGs and ExAGs is presented (Section 2.2). Sections 2.3 and 2.4 apply this framework to provide options for the SPG and ExAG, respectively. Section 2.5 reviews the options described in Gimsing et al. (2019). Finally, Section 2.6 gives some recommendations for risk managers.

2.1. Groundwater in support of and as a habitat for biodiversity

The regulatory framework for PPPs laid out in Commission Regulation (EC) No 1107/2009 and Commission Regulation (EU) No 283/2013 and 284/2013 [7]. In Commission Regulation (EC) No 1107/2009, general protection goals refer to groundwater in Article 4, approval criteria for active substances. In Section 3b of this Article it is stated that they 'shall have no immediate or delayed harmful effect on human health [...]; or on groundwater'. In Section 3e, it is further stated that they 'shall have no unacceptable effects on the environment, having particular regard to the following [...] its fate and distribution in the environment, particularly contamination of [...] groundwater [...]'. So, the Commission Regulation requires consideration of the impacts of exposure to PPPs in these

[6] In this document, 'the Panel' refers to the Panel on Plant Protection Products and their Residues (PPR-Panel).

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

environments on non-target species, on their ongoing behaviour and on biodiversity and the ecosystem, as detailed below.

In the current risk assessment for groundwater, the main ecosystem service that is implicitly addressed is 'provisioning of clean drinking water'. The importance of this ecosystem service was acknowledged in the EU Groundwater Directive,[8] which states that 'Groundwater in bodies of water used for the abstraction of drinking water or intended for such future use must be protected in such a way that deterioration in the quality of such bodies of water is avoided in order to reduce the level of purification treatment required in the production of drinking water'. Because of this, the paper of Gimsing et al. (2019) focuses mainly on the ecosystem service 'provisioning of clean drinking water', and this Statement considers this ecosystem service as well. However, when defining SPGs for groundwater, further relevant ecosystem services must be reviewed, especially those addressing groundwater as habitat for non-target species and the impact of groundwater contamination on surface water and its communities.

The preamble of the EU-Groundwater Directive also mentions the importance of groundwater for groundwater-dependent aquatic and terrestrial ecosystems. Annex 1 of the Directive states that 'where, for a given body of groundwater, it is considered that the groundwater quality standards[9] could result in failure to achieve the environmental objectives specified in Article 4 of Directive 2000/60/EC for associated bodies of surface water, or in any significant diminution of the ecological or chemical quality of such bodies, or in any significant damage to terrestrial ecosystems which depend directly on the body of groundwater, more stringent threshold values need to be established'. Article 3 of the Directive describes the consequence that 'groundwater quality criteria shall *inter alia* take into account human toxicology and ecotoxicology knowledge'. The recent Proposal for a Directive amending the Water Framework Directive, the Groundwater Directive and the Environmental Quality Standards Directive[10] states that a similar way compared to surface water 'Chemical pollution of surface and groundwater poses a threat to the aquatic environment, with effects such as acute and chronic toxicity in aquatic organisms, accumulation of pollutants in the ecosystem and loss of habitats and biodiversity, as well as to human health. Setting environmental quality standards helps to implement the zero-pollution ambition for a toxic-free environment'. So, the European Commission obviously intends to protect the aquatic environment in groundwater in a similar way compared to surface water.

Also in the literature, it has been recognised that groundwater is a critical source of biodiversity (e.g. Danielopol et al., 2000) that needs to be protected against contamination (e.g. Sket, 1999a,b). More recent studies focussed on the effect of pesticides on ecosystem services based on microbial activities. Contamination of groundwater may permanently eradicate unique groundwater microbial communities due to their inability to recolonise any affected habitats (Di Lorenzo et al., 2014, 2015). Michel et al. (2021) demonstrated side effects of pesticides on microbial denitrification. In addition, if key species and ecological functions are affected, then purification processes which are important for the ecosystem service 'providing clean drinking water', might be disturbed. Also, NRC (2000) and EFSA PPR Panel (2017) highlighted the importance of microbial communities for natural attenuation. The same can be stated for the very diverse groundwater fauna (the so-called stygofauna with more than 2000 species), dominated by arthropods and contributing to several ecosystem services (e.g. EFSA PPR Panel, 2015; Manenti et al., 2021). A recent work has reviewed the impact of stressors on these groundwater species (Castano-Sanchez et al., 2020). For these reasons, the European Medicines Agency (EMA, 2018) proposes that in addition to the human health risk assessment, an environmental risk assessment for groundwater ecosystems is also needed, and the EFSA FEEDAP Panel acknowledged the relevance of protecting biodiversity in groundwater in its recent guideline for feed additives (EFSA FEEDAP Panel, 2019).

An important point for discussion for performing an environmental risk assessment for groundwater is the lack of specific standard tests for groundwater organisms. However, as for surface waters, laboratory standard species can be employed as surrogate for the assessment of risks to specific organism groups (e.g. *Daphnia* for arthropods in aquatic systems) and could be considered with appropriate amendments also here. Given the importance of groundwater biodiversity, the Panel recommends developing an environmental risk assessment scheme for groundwater, considering both the protection of biodiversity and protection of key ecosystem services. A critical question here is

[8] https://eur-lex.europa.eu/eli/dir/2006/118

[9] The groundwater quality criterium for active substances in pesticides and their relevant metabolites mentioned in Annex 1 of the Groundwater Directive is 0.1 μg/L. See Section 2.3 for details.

[10] https://environment.ec.europa.eu/publications/proposal-amending-water-directives_en

whether the current and proposed guidance on risk assessment for non-target organisms (aquatic organisms and soil organisms) covers biodiversity in groundwater sufficiently well, via testing of sensitive groups.

2.2. Need for defining both specific protection goals, and exposure assessment goals

As described in Section 2.1, it is not clearly defined which ecosystem services procured by non-target organisms must be protected where and when, nor is protection of biodiversity considered. The gap of missing explicitly defined protection goals was noticed by the EFSA PPR Panel and a scientific opinion on the development of SPG-options for environmental risk assessment of pesticides was published in 2010 (EFSA PPR Panel, 2010). They focussed on the definition of SPGs for taxonomic groups or other ecological entities, so, on the organism groups to protect. In their figure 8 they visualise that both an exposure assessment and an effect assessment needs to be carried out. In 2016, a Guidance Document was published by the EFSA Scientific Committee (SC) on how to develop SPGs for environmental risk assessment at EFSA, in relation to biodiversity and ecosystem services (EFSA SC, 2016).

As mentioned above, SPGs are focussed on ecological entities. Basically, the SPG concept allows for an unambiguous definition of which ecosystem services procure by the non-target organisms need to be protected where and when. SPGs are defined in terms of (i) ecosystems services used to identify the involved entities (ii) service providing units and (iii) six dimensions of these services providing units to be protected. These six dimensions are ecological entity, attribute, magnitude of effect, temporal scale, spatial scale and degree of certainty (EFSA PPR Panel, 2010; EFSA SC, 2016). SPGs are operationalised by Effect Assessment Goals (EfAGs) and ExAGs, evaluating ecotoxicological effects for imposed exposure regimes and field exposure, respectively (Figure 1, reproduced from Adriaanse et al., 2022):

- EfAGs define, e.g. the reference taxa representing the ecological entity, the type of ecotoxicological effect endpoint to consider and toxicity endpoint quantifying it. Moreover, the uncertainty in the extrapolation to the SPG (represented by the surrogate reference tier, see e.g. EFSA PPR Panel, 2013a) is defined when calibrating the risk assessment scheme. As stated in Section 2.1, ecological and ecotoxicological considerations are currently not considered when evaluating the ecosystem service 'provisioning of clean drinking water'. For this reason, EfAGs have at this stage not been defined for the protection of groundwater but should be developed when SPGs are identified that consider protection of organisms in groundwater.
- For defining ExAGs, five questions were formulated to define so-called 'realistic worst case' conditions (EFSA PPR Panel, 2010). For groundwater protection, representatives of the international scientific community detailed the above-mentioned dimensions[11] at the EU Modelling Workshop in Vienna of 2014 (these are included in appendix 1 of Gimsing et al., 2019). In addition, they presented seven options for the ExAG for groundwater (called groundwater protection goal by Gimsing et al., 2019), starting with the most conservative option and ending with less protective options (appendix 1 of Gimsing et al., 2019).

Note that while SPGs are to be decided by risk managers, the operationalisation of the ExAG (and the EfAG, if applicable) must be done by risk assessors. Input from risk managers is, however, also needed when operationalising the ExAG, because some choices have political and socio-economic consequences, which are beyond the scope of risk assessors. So, risk communication is not only important when defining the SPG, but also when operationalisation of the ExAG (Figure 1; see further EFSA PPR Panel, 2010).

The ExAGs and EfAGs operationalise the SPGs with respect to the exposure in the area of use and the assessment of ecotoxicological effects, respectively, so they translate the SPGs into feasible and practical procedures. Notice that the SETAC Working Group mentioned in their Problem definition document (Tiktak et al., 2020) titled 'Spatially Distributed Leaching Modelling (SDLM) of Pesticides in the context of Regulation (EC) 1107/2009' that SPGs and ExAGs can be used as synonyms. However, based on the insights above, the Panel distinguishes between the two terms to avoid confusion with other environmental domains.

[11] Notice that the sixth dimension (the degree of certainty) was not considered.

Examples of well-defined ExAGs in other pesticide risk assessment areas have been described in EFSA PPR Panel (2012) for exposure of soil organisms, Boesten (2018) for exposure of aquatic organisms, EFSA PPR Panel (2018) for exposure of amphibians and reptiles, in EFSA (2022) for bees and in Adriaanse et al. (2022) for exposure of small mammals off field and non-target arthropods. Note that in this statement no ecotoxicological protection goal is evaluated, therefore EfAGs do not feature in this statement and the ExAG reflects the SPG closer than in case the SPG is geared towards protection of ecosystems or biodiversity.

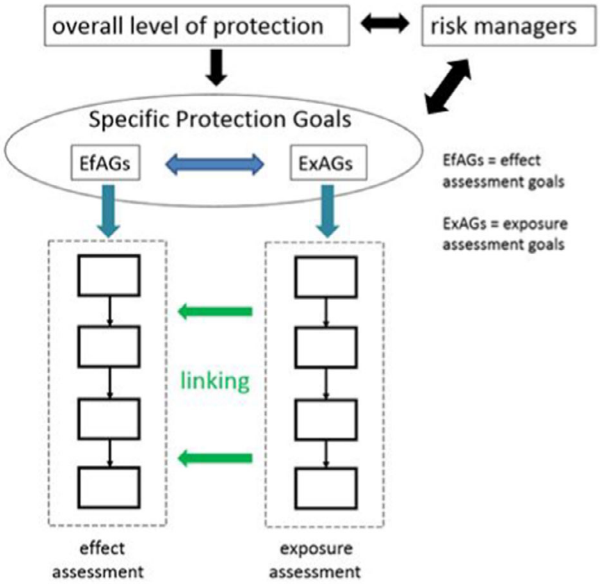

Figure 1: Schematic overview of the risk assessment of non-target organisms based on parallel tiered effect and exposure assessments that are linked to each other (adapted from Adriaanse et al., 2022)

2.3. Specific protection goals for groundwater

The need for establishing harmonised SPGs for groundwater was stated in multiple comments from stakeholders during the public consultation (Appendix D). In line with those comments, the Panel recommends that one or more SPGs should be defined by risk managers before providing dedicated regulatory guidance on how to conduct, assess and evaluate a monitoring study. Also, the definition of ExAGs by risk assessors is not possible without clearly defined SPGs (Section 2.2). The SPGs should be harmonised between the relevant European Directives, Regulations and Guidance. Based as much as possible upon existing documents, the SPG might currently read as presented in Table 1. However, the Panel recommends that further evaluation of the general protection goals is performed in future, to derive SPGs covering the other relevant ecosystems services, which groundwater provides and supports (Section 2.2).

Table 1: Overview of the SPG in the current leaching to groundwater assessment scheme according to the EFSA framework for defining SPGs (EFSA PPR Panel, 2010; EFSA SC, 2016)

Step 1. Definition of ecosystem services	Provisioning of clean drinking water.[(a)]
Step 2. Definition of the service providing unit	All groundwater that is currently used to prepare drinking water or may be used for that purpose in future.[(b)]
Step 3. Specification of the level/parameters of protection of the service providing unit based on five interrelated dimensions	
(Ecological) **Entity**	Groundwater body.[(c)]
Attribute	Concentration in pore water of the active substance, or of metabolites, degradation or reaction products of PPPs in the groundwater body.

Magnitude	Negligible effect, i.e. below maximum permissible concentrations as defined in Regulation 546/2011.
Temporal scale	Not explicitly specified, but the description of the service providing unit suggests protection over a certain time-period.
Spatial scale	Not explicitly specified. The groundwater directive suggests protection at the scale of groundwater bodies.
Degree of certainty	High level of certainty.[(d)]

(a): Notice that for the functioning of this ecosystem service, microbial and faunal communities are important (see further Section 2.1). This implicitly implies that protection of biodiversity is important as well.
(b): Definition from the Groundwater Directive (https://eur-lex.europa.eu/eli/dir/2006/118).
(c): Groundwater bodies are the managements unit under the Water Framework Directive (WFD). Groundwater bodies are subdivisions of large geographical areas of aquifers so that they can be effectively managed in order to protect the groundwater and linked surface waters.
(d): Annex II of Regulation (EC) No 1107/2009 states that the assessment should take into account the 'uncertainty of the data'. This implies that explicit consideration of uncertainty is appropriate, both by the PPR-Panel when developing the revised guidance documents, and during the assessment of specific risks.

Maximum permissible concentration

The Uniform Principles (Regulation 546/2011) state that no authorisation shall be granted if the substance or any of its relevant metabolites has a calculated or measured concentration in the groundwater of more than 0.1 μg/L or a concentration corresponding to more than 0.1 times the acceptable daily intake (ADI) in μg/kg bodyweight. In most cases, the concentration corresponding to one tenth of the ADI is higher than 0.1 μg/L, so in practice the maximum permissible concentration of active substances and its relevant metabolites in groundwater is 0.1 μg/L. A metabolite is considered 'relevant' if its toxicological properties lead to certain classifications according to the Guidance document on the assessment of the relevance of metabolites in groundwater of substances regulated under Regulation (EC) No 1107/2009[12]. The value of 0.1 μg/L (a surrogate zero) was consistent with the precautionary principle when it was originally set in 1980. It is not based on an (eco)toxicological threshold of concern, nor mixtures thereof (Dolan et al., 2013). Notice that for non-relevant metabolites there are currently no agreed values. However, the EC (2022) proposes, tiered legal threshold values for non-relevant metabolites. Based on the data available on chronic or acute effects of the non-relevant metabolites on at least one species, ecotoxicological threshold values of 0.1, 1.0, 2.5 or 5 μg/L are suggested for individual substances.

2.4. Exposure assessment goal for groundwater

Although Table 1 gives good indications for the operationalisation of the SPG, especially the temporal and spatial scale of the service providing unit are not well defined and need further operationalisation in the ExAGs. A first attempt to describe the ExAG (at that time called protection goal) for groundwater was done in 2000 when the guidance report of the FOCUS Groundwater Scenarios Workgroup was published. This report presents nine FOCUS groundwater scenarios for use in the EU pesticide registration process (FOCUS, 2000). These scenarios intended to represent realistic worst-case conditions with respect to leaching to groundwater at EU level and were aimed at an overall vulnerability covering the 90th percentile concentration in groundwater, approximated by using an 80th percentile value for soil and an 80th percentile for weather. The simulations were done for 1-m depth, which was supposed to be a conservative estimate of the groundwater concentration. In the 2014 update of the report (EC, 2014), national protection goals were suggested as follows: The annual average concentration of an active substance or its metabolites should not exceed 0.1 μg/L at 90% of the situations, taking into account both spatial variability for soil and climatic conditions, and temporal variability on a multiyear basis, in the agricultural use area of the product. There is, however, no exact definition of the spatial and temporal scales in the protection goals to be considered, nor of the depth of the assessment. Also, the EU Regulations on groundwater protection and its uniform principles (EC 1107/2009 and 546/2011) do not explicitly mention spatial and temporal scales to consider for the maximum permissible concentration in groundwater.

More recently, EFSA PPR Panel (2018) and Adriaanse et al. (2022) presented a transparent and unambiguous way to define the ExAG. They suggested six elements of the ExAG to be defined explicitly:

[12] https://food.ec.europa.eu/system/files/2021-10/pesticides_ppp_app-proc_guide_fate_metabolites-groundwtr-rev11.pdf.

- Ecotoxicological relevant concentration or ecotoxicologically relevant exposure Quantity (ERC or more general EREQ),
- Temporal dimension of ERC,
- Spatial unit, type and size,
- Statistical population of spatial units,
- Multiyear temporal statistical population of ERC values for one spatial unit,
- Desired spatiotemporal percentile of the statistical population of ERC values.

Note that for the *current* groundwater risk assessment, the ERC reduces to the **RC** (*Relevant* Concentration) as no direct impact on groundwater organisms is explicitly considered (Table 1). However, as mentioned in Section 2.1, protection of groundwater communities may require additional SPGs, and next, a separate effects assessment scheme, because it is not known if the current threshold values are sufficiently protective for groundwater communities. The six elements of the ExAG for the ecosystem service 'provisioning of clean drinking water' are proposed as described by Table 2.

The relevant concentration

Element (i) of the ExAG, the relevant concentration, is defined with the aid of the attribute of Step 3 of the SPG in Table 1. The depth at which this concentration is monitored or modelled has a clear influence on the conservativeness of the leaching assessment. Generally, it is likely that deeper depths result in lower concentrations due to e.g. dispersion and/or dissipation/degradation processes. In a regional-scale study in the Netherlands, Tiktak et al. (2005) demonstrated that the concentration at 10-m depth is generally lower than the concentration at 1-m depth. So, if the aim is to protect groundwater as a source of drinking water, the Panel recommends using a target depth above the depth where drinking water abstraction generally takes place (i.e. between 1- and 10-m depth). If the aim is to also protect associated aquatic and terrestrial ecosystems, shallower target depths might be needed (e.g. the top 1 m of the saturated zone). Notice, however, that in view of the temporal variability of the groundwater table, it will practically not always be easy to sample the top 1 m of the saturated zone, so the actual monitoring depth is a compromise between protectiveness and practicability.

Table 2: Elements of a proposed Exposure Assessment Goal (ExAG) for protection of groundwater intended for providing clean drinking water

Element	Description	Explanation
(i) Type of relevant concentration (RC)	Concentration in groundwater at a certain depth interval.	Relevant concentration, here the concentration in pore water (μg/L).
(ii) Temporal dimension of the relevant concentration	Values averaged over a certain time window. Options range from annual arithmetic mean values to instantaneous maximum values.	Annual mean values at 1-m depth below the soil surface are used in Tiers-1, -2 and -3 of FOCUS groundwater (EC, 2014)[(a)]
(iii) Spatial unit (SU), type and dimensions	Groundwater unit below a single agricultural field, represented by the depth interval[(b)] in the groundwater.	The single agricultural field was also used in EC (2014). The upper groundwater layer is expected to have conservative values of the RC.
(iv) Statistical population of spatial units	Groundwater units below all agricultural fields within a FOCUS climatic zone where a PPP is potentially being applied.	Statistical population of SUs considered in the exposure assessment. This definition is in line with the first three tiers of FOCUS groundwater (EC, 2014).
(v) Multiyear temporal statistical population of RC values for one spatial unit	The multiyear temporal population of RCs consists of a time series of RCs defined in items (i), (ii) and (iii)	The time series need to be long enough to be fit for purpose. EC (2014) uses a time series of 20 years for modelling, whereas EC (2009) proposes at least 6 years for trend analysis based on measurements.

Element	Description	Explanation
(vi) Desired spatiotemporal percentile of the statistical population of RC values	Selection of a 90th percentile is a common approach in regulatory risk assessment, but other options are possible as well.	Determines which part of the spatiotemporal population is excluded from the risk assessment.

(a): Notice that for SPGs related to non-target species providing important ecosystem services, other temporal dimensions might apply.

(b): Depth should be specified relative to the soil surface as well as relative to the groundwater table.

The temporal dimension of the relevant concentration

Element (ii), the temporal dimension of the relevant concentration describes the time window over which concentrations are averaged. Options range from annual arithmetic mean concentrations to instantaneous maximum values. Annual arithmetic mean concentrations are currently used in FOCUS groundwater (EC, 2014). When maximum concentrations or concentrations averaged over any other time window than 1 year (e.g. 3 months) are selected, it must be ensured that the FOCUS tiered approach is consistent. The Panel further notes that the variability of predicted and measured concentrations at smaller time windows (e.g. 1 day) can be relatively large. This implies that predicted or measured concentrations may temporarily exceed the standard of 0.1 μg/L, while values averaged over a certain period can be below the standard of 0.1 μg/L. Assessing groundwater concentrations based on measurements at shorter temporal scales would require short monitoring intervals.

The spatial unit

The guidance on groundwater status and trend assessment (EC, 2009) states that annual mean concentrations should be calculated for each monitoring site separately. This guidance document, however, does not exactly specify the size of the monitoring site. In regulatory risk assessment within the framework of EU Regulation 1107/2009, it is necessary to establish a link between the modelled or measured concentration and the use of a pesticide. Therefore, the Panel proposes to define the spatial unit to be the groundwater unit at a certain depth interval below a single agricultural field where the pesticide is potentially being applied. This definition is in line with for example the EFSA guidance on exposure of soil organisms (EFSA PPR Panel, 2012; EFSA, 2017) and is also implicitly used in FOCUS groundwater (EC, 2014). This definition implies that if more monitoring wells are hydrologically connected to one field, results from these wells should be averaged and be considered as results for one spatial unit. Notice that lateral flow processes and transversal dispersion can make the link between agricultural use and the groundwater concentration less straightforward, particularly if the target is deep groundwater. Demonstrating connection between agricultural use and the groundwater concentration is therefore an important issue when interpreting monitoring studies (Section 6).

Depth is an important element of the definition of the spatial unit. The spatial unit is defined by both the depth of the groundwater unit below the soil surface, and the depth interval in the upper groundwater layer of the groundwater unit. The Panel concludes that concentrations at depths ranging from 1 to 10 m below the soil surface and that originate from the shallow groundwater (1–2 m below the groundwater table), represent well, i.e. in a conservative way, the concentration in the entire groundwater body or aquifer.

The statistical population of spatial units

Element (iv), the statistical population of spatial units, is the next item to be defined. The first aspect of this population is the total area to be considered. The Groundwater Directive introduces the concept of groundwater bodies. The Irish Geological Survey gives the following definition: 'Groundwater bodies are subdivisions of large geographical areas of aquifers so that they can be effectively managed in order to protect the groundwater and linked surface waters'.[13] In many countries, groundwater bodies are linked to the River Basins defined in the Water Framework Directive. This subdivision is not relevant in the context of the approval and authorisation of pesticides. To be in line with the existing FOCUS groundwater guidance (EC, 2014), the Panel proposes to

[13] https://www.gsi.ie/en-ie/programmes-and-projects/groundwater/activities/understanding-ireland-groundwater/Pages/Groundwater-bodies.aspx

consider instead groundwater below all fields where the substance is applied each of the nine FOCUS climatic zones.

An important aspect of the statistical population of spatial units is whether the population is limited to groundwater below fields where the substance has **actually** been applied or to groundwater below all fields where the substance can **potentially** be applied (see EFSA PPR Panel, 2012). It has been regulatory practice so far to limit the area to groundwater below all fields where the substance can **potentially** be applied. The Panel recommends continuing this approach. It should be realised that while the statistical population of spatial units is formed by all groundwater spatial units located below agricultural fields in the potential area of use, in practice we only have measurements for a sub-sample of this underlying statistical population. Ideally, the descriptive statistics of the sub-sample should reflect the descriptive statistics of the underlying distribution (see further Section 6). Because of their specific properties, karstic areas should be excluded from the statistical population of spatial units. The Panel suggests these areas to be addressed in the national authorisation process.

The multiyear temporal statistical population of relevant concentrations

Element (v), the multiyear temporal statistical population of relevant concentrations, consists of a time-series of concentrations as described in element (ii) and (iii). The time-series must be long enough to be statistically relevant. FOCUS groundwater uses a time-series of at least 20 years for modelling. The guidance of groundwater status and trend assessment (EC, 2009) mentions 'if data are aggregated as annual means, the minimum required period of observations is 8 years. If data are collected with higher frequencies (half-annual or quarterly), a minimum of 6 years of data are required'. Although this is clearly less than the simulated 20 years in Tiers-1, -2 and -3 of FOCUS groundwater, the Panel recommends following the existing guidance on groundwater status and trend assessment, so a minimum of 6 or 8 years would be needed in monitoring studies.

The spatiotemporal percentile of the statistical population of relevant concentrations

Element (vi) is the spatiotemporal percentile of the statistical population of relevant concentrations. According to the Guidance on groundwater status and trend assessment (EC, BE3), good chemical status is reached if **all** monitoring points within a groundwater body comply. This comes down to using the 100th percentile.[14] For pesticide registration, the 90th-overall percentile has been used (e.g. EC, 2014). This means that exceedances of the maximum permissible concentration are accepted at a certain percentage of the groundwater spatial units located below the intended use-area of a pesticide. The choice of the spatiotemporal percentile defines the level of protection to be achieved in practice. Because this choice is a compromise between economic considerations and environmental considerations, the choice of the spatiotemporal percentile is clearly a risk management decision.

Another element of the spatiotemporal percentile is how the vulnerability is divided between space and time. FOCUS assumes that vulnerability is equally distributed between weather and space, so the 90th percentile leaching concentration is estimated by the combination of the 80th-temporal percentile of 20 annual average leaching concentrations and the 80th-spatial percentile. The Netherlands uses a 90th-spatial percentile in combination with the 50th percentile in time with the argument that it is better to protect a large area than a smaller area against peak concentrations (Van der Linden et al., 2004; Tiktak et al., 2022). EC (2014) evaluated the FOCUS vulnerability concept and concluded that there is no theoretical basis for assuming that the 90th-overall percentile can be approximated by using the 80th-spatial percentile of the 80th percentile of the yearly averaged concentrations at a certain location. They concluded that using the 90th-spatial percentile of the long-term averaged flux weighted concentrations may be a more relevant approach for a '90% vulnerability concept'. Generally stated, the procedure to determine an overall 90th probability of occurrence in time and space depends on the distribution of the relevant concentrations as defined in Table 2.

It should be noted that changing the spatiotemporal percentile may have consequences for the FOCUS groundwater tiered approach. If risk-managers would, for example, decide to use a higher number than the 90th percentile, revision of the lower tiers of the FOCUS groundwater assessment scheme might be necessary (see further Section 2.5).

[14] Notice that according to EC (2009) a certain percentage of monitoring wells can exceed the quality standards (the proposed default value is 20%) when appropriate investigations confirm that there is no deterioration in quality of waters for human consumption.

Drinking water abstraction for private uses

A monitoring survey in Suffolk County (New York) indicated already in 1982 that private wells adjacent to arable fields can become contaminated with pesticides at levels above state recommended health levels (Zaki et al., 1982). This raises the question if the risk assessment should also cover drinking water abstraction for private uses. The Panel notes that drinking water abstraction for private uses can take place adjacent to agricultural fields and at relatively shallow depths compared to drinking water abstraction for public use. Also, these private uses are generally not situated in drinking water abstraction areas, which in some countries (e.g. the Netherlands) have a higher level of protection than in areas where no drinking water abstraction takes place (Tiktak et al., 2022). Should risk managers decide that the leaching assessment should also include private uses, this should be considered in the selection of the ExAG.

2.5. Review of the exposure assessment options in Gimsing et al. (2019)

During the 7th EU Modelling Workshop in Vienna, five questions were developed to define the ExAG (see appendix 1 in Gimsing et al., 2019). These questions are comparable to the six elements of the ExAG as described in Section 2.4. However, the multiyear temporal statistical population of relevant concentrations has not been considered by Gimsing et al. (2019). The Panel considers this an important element of the ExAG.

Gimsing et al. (2019) responded to the five questions by formulating seven options for the ExAG[15] that intended to cover the full range, considered relevant by risk managers, from very strict to least strict (Table 3). Here we discuss these responses to the five questions in more general terms.

Table 3: Summary of exposure assessment options as described in Gimsing et al. (2019)

Option	Description
1	Residue concentration in the upper 10 cm of the saturated zone – including output from drains
2	Residue concentration in the upper portion of the groundwater from below treated fields but excluding groundwater shallower than 1 m below the ground surface
3	Same than option 2 but excluding areas that will never be used for drinking water production
4	Residue concentration in groundwater shallower than 10 m below-ground surface but excluding groundwater shallower than 1 m below ground surface
5	Residue concentration in groundwater deeper than 10 m below-ground surface, representing depth typical for groundwater abstraction
6	Residue concentration in raw water of an abstraction well
7	Residue concentration in raw water of an abstraction well, water not older than 50 years

Type of concentration

Gimsing et al. (2019) do not give an exact description of the type of concentration needed in the groundwater assessment (see Section 3.1 for a more detailed discussion on the type of concentration), but limit themselves to proposals on *where* concentrations should be protected, e.g. where in the saturated zone, at which depth below the soil surface, or in which type of groundwater the concentration is to be found (e.g. raw water of a drinking water pumping station without bank infiltration of a certain age). As mentioned in Section 2.4, not all options are compatible with the current FOCUS groundwater tiered approach. Particularly their first option (the concentration in the upper 10 cm of the water-saturated zone of a treated field) would be more conservative than the modelling tiers. Because the evaluation depth is interlinked with the other dimensions of the ExAG and has a clear influence on the conservativeness of the assessment, the Panel suggests that risk managers are consulted on the evaluation depth as part of the definition of the ExAG (see also Section 2.4).

[15] Gimsing et al. (2019) use the word specific protection goals. The Panel proposes using the term exposure assessment goal (see Section 2.2).

The spatial unit

Gimsing et al. (2019) define the spatial unit by the size of the area delivering input into the groundwater body (e.g. 1 m^2 of treated field or an entire treated field of 1 ha). This definition is more related to the input of pesticides into the groundwater body than to the exact description of the groundwater body itself. The Panel proposes to define clearly the groundwater body to be protected. To be in line with lower tiers in FOCUS groundwater, the Panel proposes to define the spatial unit to be the groundwater below a single agricultural field where the pesticide is potentially being applied (see further Section 2.4) and call this a *groundwater unit*. A clear definition of the spatial unit is needed to design appropriate sampling strategies.

The spatial statistical population of these units

Gimsing et al. (2019) relate the spatial statistical population of units to the potential area of use of the substance. This definition is again related to the input of pesticides into the groundwater body and not to the groundwater body itself. To be in line with the existing FOCUS groundwater guidance (EC, 2014), the Panel proposes to consider instead groundwater units below all fields where the substance is potentially applied in one of the nine FOCUS climatic zones. This definition is comparable to the definition given by Gimsing et al. (2019); however, it relates to the groundwater itself (Section 2.4).

The temporal statistical population of these units

Gimsing et al. (2019) define the temporal dimensions of the relevant concentration, e.g. daily or weekly values, or annual average values. However, the length of the time series has not been specified. The Panel recommends to clearly specify how many years the relevant concentrations should span (e.g. 6 years or 8 years, see Section 2.4).

The spatiotemporal percentile of the statistical population of relevant concentrations

Gimsing et al. (2019) present various percentiles and various ways of combining measurements in space and in time to determine the overall percentile. The choice of the spatiotemporal percentile defines the level of protection to be achieved in practice. Because this choice is a compromise between economic and environmental considerations, this is clearly a risk management decision (see further Section 2.4).

The seven options reviewed

As outlined in Section 2.4, the Panel concludes that concentrations at depths ranging from 1 to 10 m below the soil surface and that originate from the shallow groundwater (1–2 m below the groundwater table), represent well, i.e. in a conservative way, the concentration in the entire groundwater body or aquifer. It was further noted the monitoring depth is a compromise between conservativeness and practicability. A direct link between the ExAGs suggested by the Panel and the options in Gimsing et al. (2019) is not possible, because the spatial unit in Gimsing et al. (2019) has been defined in a different way. Despite this, the Panel concludes that Options 1, 3, 5, 6 and 7 of Gimsing et al. (2019) are not in line with the recommendations in Section 2.4 (see Table 3 for an overview of the options in Gimsing et al., 2019). Option 1 is very conservative and would require a complete revision of the tiered approach of FOCUS groundwater (see also Section 3). Option 5 considers groundwater at 10 m depth below the soil surface, while options 6 and 7 consider monitoring at the catchment scale. The Panel considers these options impractical, because a link between product use and measured groundwater concentrations is difficult to demonstrate. Should one of these options be chosen by risk managers, the Panel still recommends a study design with a shallower monitoring depth. This would give a conservative estimate of the groundwater concentration for options 5, 6 and 7 in Gimsing et al. (2019), but would be more practical to conduct and interpret (see also the footnote in table 1 of Gimsing et al., 2019, which states that 'studies demonstrating compliance under exposure assessment options 1, 2, 3 and 4 would usually be adequate to demonstrate compliance under options 5, 6, and 7'). The Panel considers options 3 and 5 inappropriate, because the Water Framework Directive states that all groundwater needs to be protected. Options 2 and 4 closely resemble the options suggested by the Panel. Notice, however, that to fully characterise the ExAG, all six elements of the ExAG need to be defined.

2.6. Conclusions and recommendations

- The Panel notes that groundwater is an important source of biodiversity that needs to be protected from contamination.
- Before detailed and harmonised guidance can be developed, risk managers must define which ecosystem services must be protected where and when (SPG). This definition is crucial for designing appropriate risk assessment schemes in which ecotoxicological effects can be combined with exposure measured or simulated in the field.
- The current leaching assessment scheme for groundwater aims primarily at the ecosystem service 'provisioning of clean drinking water' and does not consider impacts on organisms. However, when defining SPGs for groundwater, further relevant ecosystem services must be considered especially those addressing groundwater as habitat for non-target species and the impact of groundwater contamination on surface water and its communities.
- The exposure assessment is an integral part of any risk assessment scheme. To define which exposure should be used to evaluate the SPG for groundwater in a structured and unambiguous way, the so-called ExAG must be defined. The Panel recommends operationalising ExAGs in a transparent and unambiguous manner using the six elements of the ExAG as a guideline, considering also different SPGs. This operationalisation should preferably be done by risk assessors after consultation of risk managers. Important elements to decide upon by the risk managers are (i) the depth relative to the soil surface and relative to the groundwater layer, (ii) the time window over which concentrations can be averaged and (iii) the desired overall level of protection by selecting a spatiotemporal percentile of the statistical population of relevant concentrations.
- As water sampled in the uppermost groundwater is assumed to be on the conservative side with respect to concentrations over the entire depth (due to processes such as degradation during downward transport, or dilution with uncontaminated water from non-treated areas), the Panel concludes that concentrations at depths ranging from 1 to 10 m below the soil surface and that originate from the upper layer of the groundwater body (1–2 m below the groundwater table), represent well, i.e. in a conservative way, the concentration in the entire groundwater body or aquifer. This implies that ExAG-options 2 and 4 in Gimsing et al. (2019) closely resemble the options suggested by the Panel. Notice, however, that to fully characterise the ExAG, all six elements of the ExAG need to be defined.

3. Principle of the tiered approach and challenges for the groundwater assessment

The intention of any tiered assessment scheme is that the initial (or earlier) tiers are quick, simple and relatively cheap to undertake and allow the compounds that clearly do not cause any concern to pass (EFSA PPR Panel, 2010). The later (or higher) tiers are more complex and expensive but should provide a more realistic (often less conservative) result (EC, 2014). To maintain the consistency of the tiered approach, higher tiers should result in a less conservative value of the target exposure quantity; i.e. they should give a more realistic estimate of the target quantity of the ExAG. Figure 2 presents the tiered approach for the groundwater assessment (EC, 2014) with the groundwater monitoring tier being the highest tier.

Some stakeholders raised concerns about the position of groundwater monitoring in the FOCUS tiered approach. They argue that the lower tiers predict a concentration that is generally relevant for the unsaturated zone (and not the saturated zone of the groundwater) or a different depth. They further argue that the measured groundwater concentration depends on processes that are not considered at the lower tiers. In this section, we first discuss how the target quantity in the lower tiers relates to the groundwater concentration. Then we discuss the processes that are relevant for the concentration in groundwater, but which are not considered in the lower tiers of FOCUS groundwater.

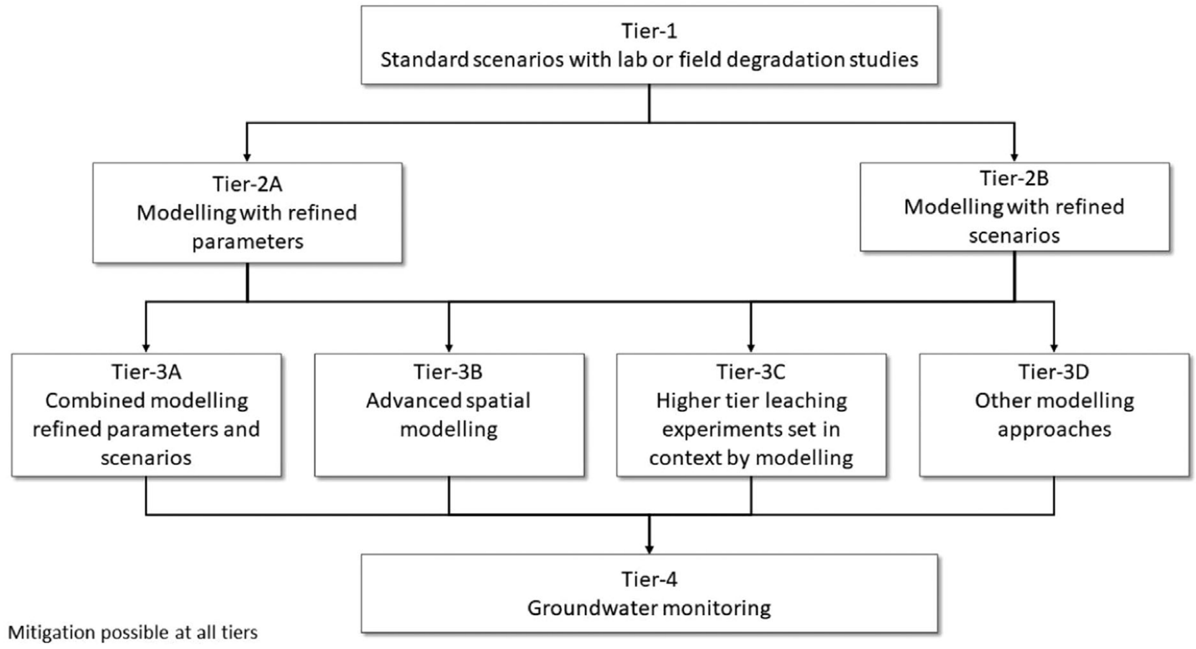

Figure 2: Tiered assessment scheme of FOCUS groundwater (EC, 2014)

3.1. The type of concentration

In a tiered approach, all tiers should evaluate the same SPG and the related ExAG, so all tiers should result in the same type of concentration (i.e. the relevant concentration). This section reviews the type of concentration in the different tiers of FOCUS groundwater and discusses consequences for the leaching assessment.

Resident vs flux concentrations

A source of confusion in groundwater leaching assessment is the difference between the so-called resident concentration and the flux concentration. The former describes all solute particles present at some point while the latter describes solute particles that have moved irrevocably past some point (Kreft and Zuber, 1978; Zhang et al., 2006).

The difference between the two types of concentrations can be several orders of magnitude (Parker and van Genuchten, 1984; Zhang et al., 2006). The discrepancy between the flux and resident concentration is dictated by the magnitude of the ratio between dispersion and mean advection (dispersivity) and mechanisms of solute retention. The discrepancy increases with increasing dispersivity and with the existence of preferential flow pathways and/or diffusion-limited immobile zones (Zhang et al., 2006), especially for short transport times or short transport distances (which is the case when monitoring wells are installed in the shallow groundwater). Deviations between flux concentrations and resident concentrations in the mobile phase diminish over time in porous systems consisting of mobile and immobile zones. This implies that deviations between flux concentrations and resident concentrations in the mobile phase will be less pronounced with increasing distance between the point of application and the point of observation, i.e. for increasing transport times. For groundwater monitoring wells, this means that deviations between flux and resident concentrations will be smaller with increasing filter depth or increasing distance between the points of application and extraction. Also, the importance of preferential transport, inherently advection-dominated, will be less important at larger transport times or transport distances (e.g. Zhang et al., 2006). Conversely, the difference in type of concentration can be decisive for the interpretation for short transport times or transport distances.

Type of concentration in lower tiers compared to Tier-4

In the lower tiers of FOCUS groundwater, the target quantity is a flux concentration. This concentration is calculated by dividing the annual mass transported over the 1-m depth horizontal plane in soil by the annual water volume transported across the same plane. Which quantity is measured at Tier-4 depends on the sampling device (Appendix A). Solution samplers in groundwater observation wells operated by suction are most appropriately viewed as (local) flux concentrations (e.g. Parker and van Genuchten, 1984). This is because water from the larger pores or from regions with higher hydraulic conductivity will contribute more to the sample than water from smaller pores or from regions with lower hydraulic conductivity. Thus, water from areas with higher flow velocities are preferably, i.e. flux proportional, sampled. Note that the principle of collecting groundwater samples from monitoring wells is comparable to the way groundwater abstraction wells for drinking water operate.

Temporal scale for averaging concentrations

Although the type of concentration appears to be the same, the time scales for averaging differ between the lower tiers (year or the 80th percentile of a 20-year period) and the groundwater monitoring samples at Tier-4 (hours). It is inherent that single daily-averaged concentrations can be higher than the yearly-averaged concentration, thus invalidating the tiered approach. Figure 3 gives an impression of the variability of the time scales used for averaging the flux concentrations at a daily basis (red solid line), a yearly basis (black solid line), and the 80th percentile of a 20-year period (black dashed line) at the reference depth of 1 m for the FOCUS Hamburg scenario.

To be consistent with the *current* tiered approach, the sampling strategy of groundwater monitoring wells and the averaging method should reflect and be representative for the annual-averaged concentration. This sampling strategy should be continued for several (6–8) years. Notice that different averaging procedures could be applied as well (Section 2.4). When maximum concentrations or concentrations averaged over any other time window than 1 year (e.g. 3 months) are selected, it must be ensured that the FOCUS tiered approach is consistent.

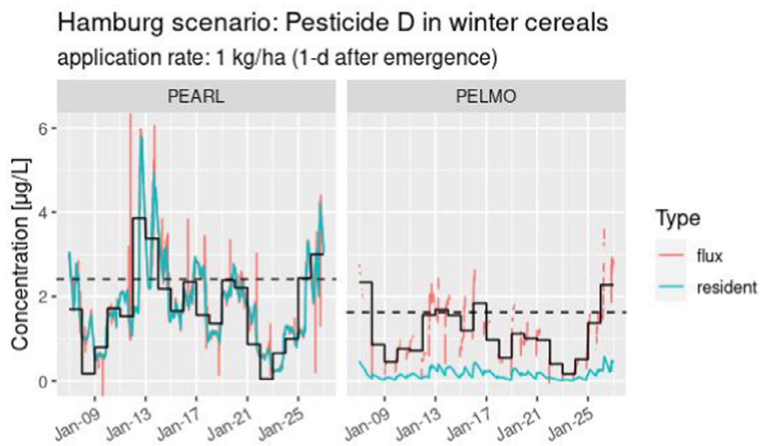

Figure 3: Comparison between the daily flux concentration and resident concentration in the liquid phase of pesticide D at 1-m depth in the Hamburg scenario calculated by PEARL v4.4.4 and PELMO v5.5.3 (see EC, 2014 for a general description of these models). Pesticide D was applied annually in winter cereals 1 day after first emergence at a rate of 1 kg/ha. The first 6 years are excluded (warming up phase). The solid black line represents the annual averaged flux concentration, and the dashed black line is the 80th percentile of the 20 annually averaged flux concentrations. Note that the flux concentration is not defined in case the water flux density equals zero, i.e. when water is stagnant, as is often the case for the PELMO model

Area-averaged vs local flux concentration

The averaging procedure differs between the lower tiers and the monitoring tier. An area-averaged flux concentration is implicitly assumed in the lower tiers (the groundwater spatial unit is the groundwater below an agricultural field of 1 ha), whereas samples from monitoring wells represent local flux concentrations. Due to spatial variability, single local concentrations (i.e. a concentration at a single point in time and space) can be higher than the average flux concentration for the entire groundwater spatial unit. Ideally, several monitoring wells should be installed to assess a spatially averaged concentration (e.g. Vanderborght and Vereecken, 2001). This, however, is seldomly the case. If multiple wells are installed under one field, the Panel recommends averaging measurements from those wells to get a better estimate of the average concentration for a spatial groundwater unit. To avoid false negatives, this should, however, only be done for wells that are connected to the field (see further Section 6).

3.2. Processes that may affect the conservativeness of Tier-4

As stated above, higher tiers should give a more realistic estimate of the groundwater concentration than lower tiers. In principle, well performed monitoring studies meet this requirement, because the effect of processes that are not modelled at Tiers-1, -2 and -3 are implicitly seen in the observations. The implicit incorporation of these processes at Tier-4 may, however, make Tier-4 either more conservative, or less conservative than the lower tiers. Processes that result in lower concentrations are for example hydrodynamic dispersion, diffusion and degradation in the saturated zone (Tiktak et al., 2005). Herrmann and Sur (2021) demonstrated that the concentration of the metabolite N,N-dimethylsulfamide (DMS) decreased continuously and consistently when the substance was followed from its formation in soil below a treated field, to its leaching into the aquifer, and within the aquifer down to the raw water collection site. Preferential flow is one of the known processes that could make the assessment at Tier-4 more conservative. It is therefore difficult to guarantee that Tier-4 is *always* less conservative than the modelling tiers.

Below, we consider three important known processes that may affect the conservativeness of groundwater monitoring (Tier-4): (i) losses by lateral flows of drainage, surface flow and interflow, (ii) the occurrence of preferential flow and (iii) canopy processes (see also Tiktak et al., 2020).

Lateral flow processes

The purely one-dimensional FOCUS leaching models do not include the loss of pesticides via surface runoff, interflow and subsurface (artificial) drainage. As the lower tier leaching assessments should represent realistic worst-case conditions, ignoring these loss processes is not an issue. This may be different for Tier-4 monitoring. In areas where lateral flow processes such as surface runoff, drainage or interflow are important, a relatively large fraction of the applied mass might be removed. The effect of this on the groundwater concentration is, however, not known. If water and substance mass are removed in equal proportions, the effect on the resulting groundwater concentration might be small. On the contrary, artificially drained soils are often also soils where macropore flow is relevant. So, it cannot be stated *a priori* that the groundwater concentration is low in areas where lateral flow processes are important. A modelling study with a spatially distributed pesticide leaching model that includes both preferential flow and these lateral flow processes could provide more insight into this question. Examples of such models are the preferential flow version of the Dutch GeoPEARL model (Tiktak et al., 2012a) and the Proziris (Burns et al., 2015) tool for the Northern European countries.

Preferential flow

Preferential flow is a general term that covers all types of non-chromatographic flow, including mechanisms such as macropore flow, funnel flow and fingered flow. There is scientific consensus that these processes occur under field conditions (Flury, 1996; Villholth et al., 1998; Scorza Júnior et al., 2007; Sanders et al., 2012; Rosenbom et al., 2015). From a pesticide leaching point of view, particularly macropore flow is interesting, because this provides a mechanism for bypassing the topsoil where a large fraction of pesticide may be removed by degradation under chromatographic flow conditions. Macropore flow occurs mainly in structured (clay) soils, although transport through biopores may be relevant in both sandy and clayey soils (Lindahl et al., 2009). Notice, however, that tillage practices may disturb existing macropores, reducing the effect of macropore flow in regular agricultural practice (Petersen et al., 2001; Ruqin et al., 2013). Shrinking and swelling of certain clay soils may result in closure of macropores early in the season, particularly in soils containing interlayered clay minerals such as smectites and montmorillonites (Bronswijk and Evers-Vermeer, 1990; van der Salm et al., 2012; Sanders et al., 2012). Except in one Tier-1 FOCUS scenario (Châteaudun), macropore flow has not been considered in the lower tiers of the FOCUS groundwater model calculations. It was demonstrated in several studies that pesticide transport can be well simulated using the convection–dispersion equation in light textured soils (see Vanclooster et al., 2000 for an overview), but that ignoring preferential flow in structured soils may result in underestimation of the mass leached (Vanclooster et al., 2000; Tiktak et al., 2012b; Kupfersberger et al., 2017), particularly for substances with a high K_{om}-value (Larson and Jarvis, 2000; Scorza Júnior and Boesten, 2005). So, the (implicit) inclusion of macropore flow in Tier-4 may in structured soils *increase* the conservativeness of this highest tier, instead of decreasing it, and so, the logic of the tiered approach could be breached.

Although there is consensus that macropore flow is important for correctly describing the transport of pesticides through the upper layers of structured soils, there is scientific evidence that many macropores end at depths near or above the saturated zone. Van Stiphout et al. (1987) demonstrated that water flowing in continuous macropores infiltrated into the subsoil at depths between 60 and 130 cm. This subsoil infiltration was called 'internal catchment'. Later, several authors confirmed that both biopores and shrinkage cracks end near the saturated zone, confirming the importance of this internal catchment for correctly describing pesticide transport in structured soils (Scorza Júnior et al., 2004; Lindahl et al., 2009; Van Schaik et al., 2010; Tiktak et al., 2012b). Whether macropore flow also leads to higher vulnerability of the underlying aquifer depends also on the thickness of the clay layer covering the aquifer. If the clay layer extends to larger depths, the macropores will end in the clay layer itself. As a result, the clay layer will act as an aquitard, and the underlying (permeable) aquifer will be invulnerable to pesticide leaching. On the contrary, if cracks and fissures extend to larger depths (e.g. karstic areas), the underlying aquifer may be highly vulnerable. It is yet unclear how often macropore flow leads to high vulnerability of the underlying aquifer and hence to an inconsistent tiered approach. An important obstacle for this is the lack of data describing the geohydrological situation of deeper soil layer at EU level.

Colloid-facilitated transport

In the public consultation, one of the Member States addresses the issue of possible transport of strongly sorbing pesticides by solids. There are indeed indications that under conditions of preferential flow, transport of strongly sorbing pesticides can be colloid facilitated (e.g. De Jonge et al., 2000; Villholth et al., 2000). In a study on radionuclide transport, de Weerd and Leijnse (1997) demonstrated that transport of radionuclides in groundwater can be enhanced or retarded because of the presence of colloids. However, all cited studies are limited to shallow depths (soil columns and drainage water), so the consequence of colloid-facilitated transport to deeper aquifers is not known. Also, models describing colloid-facilitated transport are highly complex (see e.g. de Weerd and Leijnse, 1997) and not yet operational for regulatory purposes. For these reasons, the Panel considers including this process into regulatory models not feasible for the time being.

Canopy processes

In the FOCUS groundwater scenarios, the application rate is reduced by the fraction that is intercepted by the crop and the reduced fraction is assumed to be applied to the soil (EC, 2014). However, the EFSA PPR Panel (2012) considered there is insufficient evidence to ignore wash-off from the crop canopy and thus EFSA (2017) recommended to include the effect of both dissipation at the crop canopy and foliar wash-off when the substance is applied to the crop canopy. Exclusion of foliar wash-off in lower tier FOCUS groundwater assessments may lead to underestimation of the predicted concentration in groundwater, while it occurs and thus might be reflected in sampled groundwater at monitoring sites. This could potentially make Tier-4 more conservative than the lower tiers of the groundwater leaching assessment. Because an agreed methodology to introduce foliar wash-off into the models exists (EFSA, 2017), the Panel recommends introducing foliar wash-off also into the lower tiers of the FOCUS groundwater assessment scheme.

Degradation, sorption and hydrodynamic dispersion in the saturated zone

These three processes are not considered by default in the lower tiers of FOCUS groundwater. Ignoring these processes leads to a conservative estimate of the groundwater concentration.

The FOCUS models assume no degradation in the zone below 1-m depth. There is evidence, however, that this process may occur in certain cases (see section 7.1.3 of EC, 2014 for examples). If sufficient evidence is available, degradation in the saturated zone can be included at Tier-3a of the leaching assessment. Commission Regulation (EU) No 283/2013 on data requirements does not describe harmonised requirements for degradation studies in the saturated zone.[16] The revised Dutch decision tree on groundwater leaching (appendix 2 in Tiktak et al., 2022) considers a simplified computation procedure to deal with degradation in the saturated zone. This Dutch guidance states that the applicant can perform degradation studies in materials from the saturated zone between 1-

[16] Article 7.2.3 on degradation in the saturated zone states that the type and conditions of the study to be performed shall be discussed with the national competent authorities.

and 10-m depth and show that under all chemical conditions, oxic to methanogenic, degradation is fast enough to reduce the concentration below the level of 0.1 μg/L at a depth of 10 m.[17] The so-obtained degradation coefficients are introduced into a simple calculation procedure, which assumes first-order degradation kinetics and a travel time of 4 years (corresponding to a precipitation surplus of 250 mm per year). The Panel considers this approach acceptable for Tier-3a assessments. The Panel notes that calculation procedure must be slightly adapted. First, the precipitation surplus must be in line with the precipitation surplus in the lower tier model calculations. Second, the depth must be in line with the ExAG set by risk managers. Third, the initial concentration, c_0, should be set to the target concentration in respective FOCUS scenarios or to the target concentration in higher tier modelling assessments. And fourth, dedicated technical guidance must be developed how to perform these degradation studies (e.g. the selection of representative soils).

Sorption is another process that may lower the concentration in pore water. By default, the FOCUS models do not consider sorption below 1-m depth, because the organic matter content is set to zero at larger depths.

In a nationwide modelling study in the Netherlands by Tiktak et al. (2005), it was shown that due to physical dispersion, the concentrations of pesticides were reduced. The main process here was longitudinal dispersion, and to a lesser extent also transversal dispersion.

3.3. Conclusions and recommendations

- The Panel concludes that well-performed monitoring studies (see Section 6 for requirements of such studies) provide for a more realistic exposure assessment than the lower tiers. However, it cannot be guaranteed that Tier-4 is *always* less conservative than the modelling tiers. Because monitoring studies provide more realistic exposure assessments, the Panel notes that Tier-4 studies can overrule results from lower tier studies, if they fulfil all quality criteria (exact criteria to be set in later guidance).
- Preferential flow is one of the known processes that could make the assessment at Tier-4 more conservative. The Panel recommends investigating if it is possible to include a harmonised description of preferential flow in the leaching models that are currently used in the leaching assessment and to investigate the consequences for the groundwater concentration in the aquifer. Ultimately, preferential flow could be incorporated into the lower tiers of FOCUS groundwater.
- Ignoring wash-off from crops in the lower tiers of FOCUS groundwater can also lead to an inconstant tiered approach. In contrast to preferential flow, this process can be easily incorporated in the current leaching models, because an agreed and harmonised description of this process is already available in the FOCUS groundwater models PEARL and PELMO.

4. Vulnerability assessment for site selection and context setting

As described in Chapter 4 of Gimsing et al. (2019), a central question in the design and interpretation of groundwater monitoring studies is that of groundwater vulnerability. Gimsing et al. (2019) define vulnerability in the context of groundwater monitoring as 'a measure of the potential for a substance applied at or near the soil surface during normal agricultural use to appear at *relatively* high concentrations in the groundwater'. Implicit in this definition is some means of comparing conditions in one location with another. In regulatory assessments, an important question is if the selected monitoring sites are sufficiently vulnerable, given the selected ExAG. To find sites that represent realistic worst-case conditions, groundwater vulnerability must be modelled for the entire use area of a pesticide. This chapter reviews the approaches suggested and gives recommendations for guidance development.

4.1. Consistency with the proposed exposure assessment goals

Gimsing et al. (2019) describe seven options for ExAGs ranging from extremely conservative to pragmatic. The Panel reviewed these options (see further Chapter 2) and concluded that not all options would be compatible with the ExAG used in the lower tiers of FOCUS groundwater (EC, 2014). The Panel further concluded that some ExAGs would be unpractical because following those it would

[17] Notice that the depth of 10 m is in line with the Dutch Exposure Assessment Goal. A different depth can be chosen as part of the exposure assessment goal.

be difficult to relate pesticide use to monitored concentrations, which is essential in regulatory assessments. Based on these considerations, the Panel concludes that concentrations at depths ranging from 1 to 10 m below the soil surface and that originate from the upper layer of the groundwater body, represent well, i.e. in a conservative way, the concentration in the entire groundwater body or aquifer (Section 2).

Depth

The suggested vulnerability mapping approaches typically predict groundwater vulnerability at a depth of 1 m below the soil surface (Gimsing et al., 2019). The concentration simulated at 1-m depth is considered a conservative estimate of the concentration in deeper layers (Tiktak et al., 2020). The depth of 1 m is a practical choice: data for deeper soil layers is generally not available at EU-level. Furthermore, the depth of 1 m is compatible with the lower tiers in FOCUS groundwater. If risk managers would decide that the ExAG would aim at protecting groundwater at deeper depths, the suggested vulnerability mapping approaches aiming at leaching at 1-m depth are not sufficient. First, it cannot be guaranteed that groundwater recharged from a field treated with a plant protection product, is hydrologically connected to the groundwater well. Second, the travel time to deeper groundwater monitoring wells may be longer than the actual time of use of a product. In this case, the well would be unsuitable to demonstrate a safe use. Third, the texture and organic matter content in deeper soil layers may be different from the soil properties of the top first meter. The consequence is that the vulnerability of deeper groundwater layers can be different from the vulnerability at 1-m depth. These three items imply that additional work needs to be done for the interpretation of groundwater monitoring studies. Gimsing et al. (2019) provides several examples, but guidance on how to perform this additional work is missing. For a harmonised risk assessment, the Panel considers further guidance development on this additional work necessary (see also Section 6).

Endpoint for the vulnerability assessment

The models currently used for vulnerability assessment in Gimsing et al. (2019) have a target depth of 1 m below the soil surface, thus focusing on leaching vulnerability rather than aquifer vulnerability. The Panel notes that vulnerability of the groundwater aquifer may differ from the groundwater vulnerability at 1-m depth. This gives uncertainty in the vulnerability assessment, which must be considered when selecting the monitoring sites and interpreting results at a later stage.

Different metrics (model outputs) can be used to assess leaching vulnerability. In one of the illustrative cases, annual mass fluxes were used as an indicator of groundwater vulnerability in the context of monitoring studies (Syngenta, 2014). They argue that flux weighted concentrations may be less suitable as indicator of groundwater vulnerability in the context of groundwater monitoring studies because high annual leachate concentrations can be associated with minimal substance mass fluxes if the water volume flux is minimal. Since mixing of these very low mass fluxes in the upper few centimetres of the groundwater already led to a huge decrease in the concentration, it can be assumed that leachate concentrations that are associated with a relevant volume flux are of higher importance for the groundwater quality. Similarly, the selection of sites in statistical monitoring studies may require an aggregation of potential leachate to a larger scale. Mass fluxes can be meaningfully added or averaged whereas it is more difficult to do this with concentrations. So, the question is: What is a better indicator of groundwater vulnerability, annual averaged flux concentration or annual mass flux?

In the lower tiers, the regulatory groundwater threshold value is the flux-weighted annual average concentration across a unit cross-sectional area perpendicular to the main direction of flow at 1-m depth below the soil surface based on chromatographic flow (the relevant concentration). This concentration is defined as the ratio of the annual mass flux and the annual water flux. It is plausible that the extent of groundwater deterioration by pesticides is positively correlated to the pesticide mass that enters the system. Thus, the annual flux-weighted concentration and mass fluxes come into question as an indicator of groundwater vulnerability.

For illustrative purposes, we define two hypothetical agricultural fields A and B with different soil properties, reflecting a sandy and a loamy texture, respectively. The annual water fluxes (net groundwater recharge) are 0.5 m for field A and 0.05 m for field B. The difference is due to the more pronounced capillary barrier effect in field A. We further assume that the microbial activity is lower in field A compared to field B. This results in a higher annual mass flux of the pesticide in field A (5 mg/m^2) compared to field B (0.5 mg/m^2). Although more pesticide mass entered the groundwater from field A (5 mg) compared to field B (0.5 mg), the flux-weighted annual average concentrations are identical (10 mg/m^3). Thus, the flux-weighted annual average flux concentration is not a good indicator of the mass entering the groundwater.

Therefore, the evaluation of the impact of pesticide applications from agricultural fields on the quality of groundwater resource requires the quantification of the mass flux to the groundwater and through the groundwater to the groundwater monitoring wells (e.g. Jamin et al., 2012). This should, however, not be accompanied by excessive water fluxes, which would result in low concentrations. This special case did not occur in numerical simulations based on a limited number of combinations of soils and weather defined by the FOCUS scenarios (see Appendix B). The percentile of the annual mass flux in time should be defined for the vulnerability indicator in a guidance document. An 80th percentile in time would be in line with the lower tiers of FOCUS groundwater.

With the proposal of annual mass flux as target quantity for the indicator of leaching vulnerability, there is a difference with the target quantity used in the lower tiers, which is an annual flux weighted concentration. At first glance, this suggests an inconsistency in the tiered approach. However, it should be noted that the annual mass flux in a vulnerability analysis is an indicator of leaching vulnerability and not a regulatory endpoint. Therefore, there is no need for a direct comparison between the target quantities used for the lower tiers and the indicator of a vulnerability assessment. Furthermore, numerical simulations showed a strong correlation between the annual mass flux and the annual flux averaged concentration (see Appendix B).

4.2. Intrinsic and specific vulnerability

Gimsing et al. (2019) describe broadly two factors that determine the overall vulnerability of groundwater to pesticide leaching. *Intrinsic vulnerability* is determined by the natural factors that affect the leaching of a pesticide. Climatic conditions, soil type and the geohydrological situation are the most important factors. *Specific vulnerability* is affected by the properties of the chemical substance and the actual use of the substance (I.e. use intensity and agricultural practices). Gimsing et al. (2019) argue that groundwater vulnerability is affected by a complex combination of processes. The implication is that some sort of process-based modelling is needed to integrate these complex interactions. Alternatively, statistical (meta)model approaches could be used, if these capture the most important processes and give the same spatial pattern of vulnerability as the original model (see further Section 4.3).

It was demonstrated in earlier studies that substance properties play an important role in the vulnerability of groundwater (e.g. Boesten and van der Linden, 1991). Heuvelink et al. (2010) and Van den Berg et al. (2012) showed that organic matter is the most important soil parameter. Tiktak et al. (2003) demonstrated that leaching of mobile substances correlates well with soil texture whereas leaching of substances with a higher K_{om}–value correlates well with soil organic matter. So different spatial patterns and hence a different worst-case situation is predicted. This is particularly the case when substance properties depend on other soil properties such as pH. In that case, even opposite spatial patterns can be predicted. The consequence is that it is not possible to derive worst-case conditions without mapping the *specific* vulnerability of groundwater.

In monitoring studies, usually not only the active substance is measured, but also one or more metabolites. Specific groundwater vulnerability can be different for each substance. Theoretically, different sites would need to be selected for each substance. This is, however, not very practical. The Panel therefore recommends selecting the site based on the properties of the substance with the most important concentration exceedances in the lower tier assessments (which is not necessarily the active substance). The actual vulnerability for the other compounds can then be assessed in a context-setting procedure afterwards.

4.3. Suitable models for vulnerability mapping

As described in Gimsing et al. (2019), several vulnerability mapping approaches exist. These approaches can be subdivided into three classes (EC, 2014):

- index models,
- process-based models,
- statistical models, including metamodels of process-based models.

Index models are basically spatial overlays of maps with the spatial distribution of groundwater vulnerability indicators such as organic matter content, clay content and precipitation. The total vulnerability is calculated according to logical or arithmetic rules. Sometimes weighting factors are applied to account for the different sensitivity of pesticide leaching to groundwater leaching indicators (e.g. EC, 2014). Index models are easy to use but cannot describe more complex interactions between different parameters. This can be done using spatially distributed process-based models. Typically, spatially distributed process-based models are based on leaching models, which are parameterised for many scenarios that represent specific locations (see further Gimsing et al., 2019). The Panel considers using the same models as in the lower tiers of FOCUS groundwater to be an advantage. There are, however, also trade-offs. First process-based models have high computational demands. Second, these models require detailed information on the environmental conditions, such as soil, climate and crop data. These problems can be overcome by using statistical metamodel approaches, which is the third category described by Gimsing et al. (2019). These models are based on correlations between pedoclimatic variables and the predicted leaching mass or concentration. Like index models, statistical metamodels are easy to use. Their use is, however, limited to areas for which the metamodel has been calibrated.

Recommended model types

The Panel considers both process-oriented models and statistical metamodels suitable to map specific groundwater vulnerability. In principle any of the models listed in Gimsing et al. (2019) could be used for this purpose. However, to ensure compatibility of the tiered approach, the Panel recommends that the model descriptions and parameterisations in EC (2014) serve as a benchmark. Because further harmonisation of, for example aboveground processes and crop development has been performed during the development of the EFSA soil guidance, it is essential that EFSA (2017) is also considered.[18] Notice that this implies that when statistical metamodels are to be used, these should be based on models that are based on EC (2014) and EFSA (2017). To facilitate efficient use of the models for regulatory purposes, the Panel recommends the development of user-friendly software tools. Version control (for example under the umbrella of FOCUS version control) would be recommended as well. Guidance in the Good Modelling Practice Opinion (EFSA PPR Panel, 2014) could help to address uncertainties in a systematic way.

Preferential flow

One disadvantage of the FOCUS groundwater models is that they do not consider preferential flow and lateral flow processes (see further Section 3). If used in a spatial context, these models will generally predict the highest leaching concentration in light textured soils with a low organic matter content. If preferential flow is incorporated into these models, structured (clay) soils might become more vulnerable to pesticide leaching (Jarvis et al., 2009; Rosenbom et al., 2009). Whether the underlying aquifer becomes also vulnerable depends also on thickness of the clay layer and whether a drainage system is present that could divert mass towards the surface water instead of the groundwater. To gain more insight into the impact of macropore flow and drainage on groundwater vulnerability, the Panel recommends performing a study with a model that incorporates these two processes (see further Section 3). Comparison with monitoring results from monitoring programmes (Section 7) would improve confidence in these models. Based on this study, the need to update the FOCUS groundwater models and the lower tiers of FOCUS groundwater could be investigated. If this work has not been done, the Panel considers application of the current models for vulnerability mapping acceptable.

Geodata for modelling

Models assessing the fate of pesticides in the environment are inherently data intense and become more so, when applied at EU level. Irrespective of the model, the required data set can be divided in a limited number of categories, including soil data, climate and weather data, and crop and management data. Gimsing et al. (2019) give a general overview of data sources, but do not give recommendations for specific data sets that should be used. The Panel considers the availability of harmonised and well documented data sets critical towards successful and reproducible application of

[18] Notice that PEARL and PELMO have gone through this harmonisation process. If other models (for example MACRO or PRZM) are used, it is essential that they are benchmarked against these two models and that process descriptions and parameterisation should be harmonised where possible.

spatial models in a regulatory framework. The availability of a harmonised data set will also increase consistency of the simulations and makes it easier for risk assessors and regulators to review the results from such simulations.

Gimsing et al. (2019) argue that data sets on cropping and climate are evolving rapidly, and that recent data must be used. The Panel agrees to this and therefore recommends regular updates of the data set and version control, e.g. under the umbrella of the FOCUS version control group. The data incorporated in the data set should cover the entire use area, be well documented, be open source and be derived in a scientifically sound and reproducible manner. More detailed quality criteria are described in Chapter 3 of the SDLM-problem definition document (Tiktak et al., 2020). A point not considered in the SDLM-problem definition document is a possible validation of the maps against point data such as data from the LUCAS (Land Use/Cover Area frame statistical Survey Soil) database (Orgiazzi et al., 2018). The Panel further notes that soil properties should be relevant for the area of use of a substance. Because the organic matter content is by far the most important soil parameter for the leaching assessment (Van den Berg et al., 2012; Heuvelink et al., 2010), using an organic matter map specific for arable soils would be necessary. A promising map would be the organic matter map based on the SoilGrids database described in De Sousa et al. (2022). Notice that the Joint Research Centre (JRC) is also developing an organic matter map that distinguishes between arable and permanent land-uses; however, this map was not available for review when this Statement was written.

4.4. Use of vulnerability mapping for site selection and context setting

Applicants must demonstrate that the selected sites represent realistic worst-case conditions as specified in the ExAG. For Tier-4, monitoring sites can be selected through one of the following two procedures (see also EFSA, 2017):

- Random sampling in combination with appropriate statistical assessment of the spatiotemporal percentile in the ExAG.
- Modelling with a spatially distributed pesticide leaching model combined with geostatistical analysis that enables a more targeted sampling strategy.

It is to be expected that hundreds of sites will be needed to assess a spatiotemporal percentile with sufficient accuracy based on measurements alone. The alternative proposed by Gimsing et al. (2019) is to use one of the leaching models to find the appropriate locations for monitoring studies. Given the workload involved with random sampling, the Panel considers the combination of monitoring and modelling the only practical way forward.

The models are foreseen to be used for two purposes, i.e.

- Identification of potential monitoring sites. In this step, modelling is used to identify areas that are vulnerable to pesticide leaching and where monitoring wells could be installed (section 4.3.3 in Gimsing et al., 2019).
- Setting monitoring studies into a larger spatial context (section 4.3.4 in Gimsing et al., 2019). In this step, site-specific modelling is confronted with results from the spatially distributed model to demonstrate that the selected monitoring sites really represent worst-case conditions.

4.4.1. Identification of potential monitoring sites

As mentioned in Gimsing et al. (2019), there is currently no generally available data set that can be used to identify potential monitoring sites *at a field-level* across the EU. It is unlikely that one of sufficient quality will become available soon and the computational burden involved in performing a vulnerability assessment on every field in Europe is at present too large to make this a practical option. Nevertheless, models can be a helpful tool to select vulnerable *regions* in which to install groundwater monitoring wells. It must be understood that modelling provides only a starting point of site selection. So, the Panel proposes to clearly distinguish between two phases of site selection (see also the SETAC-SDLM problem definition document, Tiktak et al., 2020):

- Selection of vulnerable areas and/or regions (preselection phase);
- Selection of the actual monitoring sites within these vulnerable areas.

Preselection phase

The Panel confirms that modelling can play an important role in the preselection phase because it provides probably the only way of estimating vulnerability of groundwater to leaching of a chemical in a consistent way across all the EU and of comparing leaching vulnerability of one location to another. A boundary condition for the application of such models is the availability of user-friendly, harmonised models (including the required geodata) and guidance on how to perform the vulnerability assessment. If possible, these models should be tested using results from monitoring programmes. For guidance development, several questions need to be addressed (see also the SETAC-SDLM problem definition document):

- Which metric is most appropriate to use for vulnerability assessment from those available (e.g. concentration, mass flux etcetera)? See Sections 3 and 4.1 of this Statement for considerations.
- At which spatial and temporal scale should the vulnerability assessment be done? To reduce computation times, vulnerability mapping is often done for unique combinations of soil, weather and climate. Is this resolution sufficient to identify vulnerable areas?
- Are there any substances for which the use of currently used FOCUS models would be inappropriate? For example, substances that do not show significant correlation between leaching behaviour and different soil properties included in currently available leaching models or substances that show an extreme variability of substance properties.
- Can simple metamodels be an alternative to the currently used process-based models and what would be the conditions for their use?

Notice that groundwater vulnerability is substance specific. The Panel considers it acceptable to base the site selection on the substance with the most important concentration exceedances in the lower tier assessments. It must be demonstrated in the context setting procedure performed later, whether the selected sites are sufficiently protective for the other substances too.

Final selection of monitoring sites

Once a region has been selected using the model, monitoring sites within that region need to be selected. The first selection criterion should be that the actual site characteristics are the same as those derived from the modelling exercise. This includes for example soil and weather conditions. A second group of selection criteria is additional to those that can be derived from the vulnerability assessment. This includes such items as hydrological connectivity, travel times and product use history. These items are further considered in Section 6. Gimsing et al. (2019), note – based on experiences and comments of applicants – that there is a high rate of drop-out of potential sites because farmers might not cooperate, or use different products (see e.g. Syngenta, 2014). The Panel therefore recommends considering all fields with a vulnerability score **above** the spatiotemporal percentile in the ExAG (e.g. the 90th percentile), rather than using a narrow percentile range.

4.4.2. In-context setting of monitoring sites

The second application of vulnerability mapping associated with monitoring studies is to put monitoring data or sites in context of the entire use area within a FOCUS climatic zone. The exercise is done in two steps: First, site-specific simulations are performed with the leaching model. Second, results from these site-specific simulations are plotted at the cumulative frequency distribution for the entire use area within a FOCUS climatic zone. The site is suitable if the spatiotemporal percentile of the site is larger than the spatiotemporal percentile in the ExAG. The Panel notes that this exercise is only valid if the same leaching model is used for the vulnerability mapping and the site-specific simulations. Furthermore, the exercise is valid for the top 1 m of the soil column. The Panel considers this acceptable, because vulnerability is to a large extent determined by degradation and sorption in the topsoil. Nevertheless, a critical evaluation of hydrological connectivity and travel times is needed in addition (Section 6) and monitoring sites must be discarded if they do not meet the requirements set in Section 6. Finally, the exercise must be done for all substances under evaluation, i.e. the parent compound and their metabolites.

Site-specific data

The in-context setting procedure relies on site-specific data. The notifier must report for each monitoring site the soil properties to a depth of at least 1 m. Properties that need to be reported are organic matter or organic carbon content, bulk density, texture, pH and depth to groundwater; analytical procedures need to be reported as well. If site-specific information on soil hydrological properties is available, these need to be reported. If this information is not available, the Panel considers application of pedotransfer functions acceptable; however, all models should use the same pedotransfer functions. Climate and weather data must be measured on-site or be taken from a neighbouring weather station. In the latter case, it must be demonstrated that the weather station is representative for the monitoring site. Finally, pesticide application procedures, crop type, crop properties and crop management practices must be reported. With respect to the application procedures, timing and application rates are mandatory. Preferably, also the formulation type (e.g. wettable powder, emulsifiable concentrate or granule) and the application method (e.g. hydraulic sprayer) need to be specified as well. Monitoring sites for which insufficient supporting data is available should be excluded from the analysis. In future guidance, criteria should be clearly formulated when supporting data are sufficient.

Substance properties

To ensure consistency of the tiered approach, the same substance parameters as those used in the lower tiers of the groundwater leaching assessment should be used. This also implies that all processes that are considered in Tier-2 can be taken on board in higher tiers as well. If, for example, experimental evidence exists on hydrolysis and this process is accepted at Tier-2, it can be included in the vulnerability assessment. Specific attention is needed for substances that show sorption dependent on soil properties other than organic matter (e.g. pH or clay content). The dependence of sorption on these properties is not described in the FOCUS scenarios because pH and clay content were not selection criteria when developing these scenarios (EFSA PPR Panel, 2013b). For this reason, it has been common practice to perform calculations with two contrasting pH-values for all defined FOCUS scenarios.

So, to realistically model spatial patterns of leaching for substances showing sorption dependent on soil properties other than organic matter (e.g. pH), the dependency of sorption on these properties must be introduced. In case of pH-dependent sorption, the sigmoidal function proposed by Van der Linden et al. (2009) may be applied. Alternatively, the dependency of the sorption coefficient on soil properties can be described using mathematical rules. If this option is chosen, the notifier must provide statistical evidence that this relationship occurs (EFSA, 2017). Sections 3.6 and 3.7 in Boesten et al. (2015) and sections 3.4 and 4.4 in Holdt et al. (2021) provide additional guidance on the parameterisation of the sorption coefficient.

If additional analyses reveal that degradation of the substance is dependent on (correlated to) one or several soil parameters, this dependency should also be introduced (see paragraph 3.2.2 in Tiktak et al., 2022) to generate realistic vulnerabilities. In sections 3.3. and 4.3 of Holdt et al. (2021), it is suggested to test the influence of the soil pH on degradation rates by means of a linear relationship.

Dealing with uncertainty of substance parameters

Some stakeholders raised concerns that substance properties for a given field differ from generic ‘average’ properties derived from the above-described general rules. It would therefore not be possible to plot the vulnerability of individual fields at a distribution curve of leaching concentrations. They further commented that a higher spatiotemporal percentile should be used, because ‘there is no safety factor or conservatism left at FOCUS Tier-4’. The Panel agrees that pesticide fate parameters for sorption and degradation are location dependent. However, their variation cannot be predicted or derived from other soil properties, so these parameters are often treated as stochastic parameters (Vanderborght et al., 2011; EFSA PPR Panel, 2012; Vereecken et al., 2016).

A possible solution to deal with this uncertainty would be to use worst-case values of the degradation half-life and the sorption coefficient on organic matter. However, in earlier consultations, stakeholders confirmed that ‘average’ substance parameters should be used, and that uncertainties should be considered in the selected scenarios (EFSA PPR Panel, 2012). Following EFSA PPR Panel (2012), an alternative procedure would be to use ‘average’ substance parameters and to shift the spatiotemporal percentile in the monitoring Tier-4 to higher values (e.g. the 95th percentile). The latter would be a compensation for the wider distribution of predictions of the leaching concentration in a certain region for a certain time (Heuvelink et al., 2010; Vanderborght et al., 2011). The Panel recommends investigating if this pragmatic approach would also work for the in-context setting of monitoring sites. If such a shift is accepted, the monitoring Tier-4 might result in higher concentrations than concentrations in lower tiers, thus the consistency of the tiered approach should be investigated as well.

Using in-context setting for read-across to different zones or Member States

In pesticides regulation, different spatial scales are relevant. For the approval of active substances in the EU, the whole area of use of a pesticide in the EU is pertinent. For a substance to be approved, the 'safe use' in at least one of the FOCUS climatic zones representing a major agricultural region in the EU must be demonstrated. For the authorisation of PPPs in the Member States, the risk assessment is performed at the level of the regulatory zone (in the core assessment of the registration report) and at the level of the Member State (national addendum of the registration report). At the zonal level, an agreed concept comparable to the 'one safe use' approach does not exist. The prerequisites to grant authorisations at Member State level are diverse and Member State specific. Typically, national decisions are based on the results of a specific set of the FOCUS groundwater scenarios, and refinements with higher tier data are often possible. Some Member States have defined ExAGs that serve as the basis for national decision making.

It often occurs that monitoring is performed in part of the EU only. The Panel notes that the in-context setting procedure can be used to demonstrate a safe-use in different FOCUS zones than the zone(s) where the monitoring has been performed. For this, results from the site-specific modelling must be plotted at the cumulative frequency distribution for the entire use area of the other FOCUS climatic zone(s). The Panel acknowledges that management practices may differ between FOCUS zones; e.g. irrigation may be performed in one FOCUS zone and not in another zone. These differences must be accounted for in the spatially distributed model used for the in-context setting procedure and must be thoroughly described.

The in-context setting procedure can also be used for the authorisation process at the level of regulatory zones and Member States. However, conceptual approaches to assess groundwater vulnerability may differ between Member States, and even when the same conceptual approach is used, Member States may require the use of national data for the in-context setting procedure. Furthermore, the ExAG may differ between Member States. These differences must be accounted for in the in-context setting procedure.

4.5. Conclusions and recommendations

- It was demonstrated in several studies that substance properties play an important role in the vulnerability of groundwater to pesticide leaching. The consequence is that it is not possible to derive worst-case conditions without mapping the *specific* vulnerability of groundwater. Both process-based simulation models, and statistical metamodels could be used for this purpose.
- The Panel recommends developing harmonised models for vulnerability mapping. This should include the development of harmonised geodata. The models and data sets should be well-documented, brought under version control and preferably be accessible through a user-friendly interface.
- The Panel recommends using the annual mass flux as indicator of groundwater vulnerability.
- The models that are currently used for the leaching assessment at EU level are limited to the top 1 m of the soil column. Given the lack of data on deeper soil layers, and because groundwater vulnerability is to a large extent determined by processes occurring in the topsoil, the Panel considers this the only practical way forward. However, consequently, leaching vulnerability in areas prone to preferential flow might be underestimated with this approach.
- Although groundwater vulnerability is substance dependent, the Panel accepts to base the selection of monitoring sites on the substance with the most important concentration exceedances. Whether the selected sites represent also realistic worst-case conditions for the other substances must be demonstrated in the context setting procedure performed later.
- Given the available geodata at EU level, it is unlikely that groundwater vulnerability can be assessed at the field level. Applicants must therefore demonstrate that the actual site characteristics (e.g. soil and climate data) are comparable to those derived from the vulnerability assessment. Criteria need to be set to decide when these characteristics are sufficiently equal.
- Hydrological connection between the groundwater directly below the field and the monitoring well must be demonstrated for each selected monitoring site. Other items to be investigated are travel times and product use history (see further Section 6).

- The Panel recommends dealing with uncertainty of substance fate parameters in a pragmatic way. This could be achieved by shifting the spatiotemporal percentile in the ExAG to a higher value.
- The in-context setting procedure can be used to demonstrate a safe use in different FOCUS zones than the FOCUS zone(s) where the monitoring has been performed. The in-context procedure can also be used for the authorisation process at the national level, if requirements set by national authorities are met.

5. Data quality considerations

Section 5 on data quality considerations in Gimsing et al. (2019) is well elaborated with respect to the different steps that are crucial in conducting and evaluating a groundwater monitoring study, concerning the general study quality criteria, the installation of new or the selection of existing monitoring wells, the collection and transport of samples and sample analysis. Although the section addresses the crucial steps and gives pragmatic and reliable solutions to potential problems that can arise, some Member States expressed the need to develop a regulatory guidance framework for conducting, assessing and evaluating monitoring studies in the consultation. Future guidance needs binding procedures and clear instructions when the available information is sufficient for the use in Tier-4 risk assessment. Decision trees should also be included, when several options exist.

5.1. General issues

Good laboratory practice

Gimsing et al. (2019) describe the general study quality criteria partly vaguely in terms of that the material and methods and the scientific evaluation should be reported 'in sufficient detail' or 'all relevant data generated' should be reported. In future guidance, it should be clearly stated what information is needed to accept a targeted monitoring study.

The Panel notes that targeted groundwater monitoring studies supporting Tier-4 assessment should be conducted according to Good Laboratory Practice (GLP) quality standards. This is not always possible for public groundwater monitoring studies that are generally not conducted or sponsored by registrants. The Panel notes that available public monitoring studies need to be added to the dossier[19] and that they can be used as supporting information. However, the Panel recommends that public monitoring studies must comply with defined quality standards. These defined quality standards should be clearly specified in future guidance (Section 7).

Installation of the monitoring wells

The Panel already concluded in Section 2.6 that an ExAG for groundwater at EU level should be defined before dedicated regulatory guidance can be produced on how to conduct, assess and evaluate monitoring studies. Then, the assignment of requirements and quality criteria of the well design can be better tailored to meet the ExAG. This largely affects issues related to the target depth. For example, whereas sound knowledge of the (hydro-)geology is essential for the interpretation of monitoring studies in deeper groundwater, contamination during study performance or local/temporal land use changes might be more important for monitoring studies in shallow groundwater. These differences need to be reflected in future guidance.

A variety of well designs for collecting groundwater samples exists, as well as guidelines. The most important steps to be followed in the installation of monitoring wells are summarised in the section of the same name in Gimsing et al. (2019). The Panel wants to stress the importance of a good sealing around the casing in the unsaturated zone to avoid artificial (preferred) flow paths directly in the groundwater and a respective contamination. In future guidance, a more detailed description of the well design and the installation technique under site-specific hydrogeological conditions is needed. Clear criteria must also be given for the choice of a typical design, e.g. when can sampling lances be accepted as a method to monitor shallow groundwater instead of a monitoring well?

[19] Article 7.5 of Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market.

Selection of existing monitoring wells

Existing monitoring wells have the advantage that the installation is omitted. However, the study objective, i.e. the depth of the well screen (target depth) and study design, must comply with the ExAG. Several criteria are given in Gimsing et al. (2019) that should be considered. Also in this case, the Panel recommends formulating minimal binding criteria to accept an existing well for use in a Tier-4 risk assessment.

Collection of samples

Collection of groundwater samples is described in detail in the section of the same name in Gimsing et al. (2019), covering all relevant topics of prevention of sample contamination, possible sorption of the substances to sampling material and containers, water removal by pumps, purging, sample collection, and sample transport and storage. Note that the definition of the ExAG also invalidates groundwater monitoring studies that are not aimed at this goal, because they were conducted with a different objective and therefore have a deviating study design and sampling procedure. Gimsing et al. (2019) mention that other types of wells could be sampled using simple methods such as 'turning on the faucet'. The Panel notes that these sampling methods cannot be used for Tier-4 monitoring studies.

Gimsing et al. (2019) argue that including solids in samples must be avoided, particularly for strongly sorbing compounds. The Panel agrees that this should be avoided, because the regulatory endpoint refers to a concentration in the aquatic phase. However, if the aim would be to test if colloid-facilitated transport has occurred, analysis of the filtrate could be a valuable addition (see further Section 3).

Sample analysis

The relevant requirements for the analytical method are summarised in the section of the same name in Gimsing et al. (2019). The Panel agrees to the recommendations on analytical methods in Gimsing et al. (2019). The Panel further agrees that chemical analysis should be conducted under GLP conditions. Chapter 5 of Gimsing et al. (2019) refers to guidelines for the installation of monitoring wells, for the collection of samples and for sample analysis. Since these will inherently change over time, references to guidelines should generally be stated without revision numbers in future guidance. Furthermore, analytical method validation should be applicable with SANTE guidance documents.

5.2. Conclusions and recommendations

- The excerpt on data quality considerations is well elaborated in Gimsing et al. (2019) with respect to the different steps that are crucial in conducting and evaluating a groundwater monitoring study.
- The Panel notes that targeted groundwater monitoring studies should be conducted according to GLP quality standards. Public monitoring data are often not conducted according to GLP but can provide important supporting information if they comply to defined quality standards.

6. Study design of targeted monitoring programmes

In the public consultation, Member States expressed the need to develop a regulatory guidance framework for conducting, assessing and evaluating monitoring studies at different assessment levels (e.g. Regulatory Zones, FOCUS Zones or Member States). They demanded quality criteria to enable an overall reliability assessment for each monitoring site. Distinctive advice was requested on practical issues like the minimum and maximum field size, minimum number of sampling wells, minimum number of supporting wells, minimum depth to groundwater, minimum sampling frequency and minimum study duration. Overall, the comments made strong appeals for a harmonised study design and binding procedures to facilitate regulatory acceptance of the tiered groundwater risk assessment.

Gimsing et al. (2019) highlight many practical aspects that need to be considered during the field phase of monitoring studies. Especially in Chapter 5, technical requirements for well installation, collection of samples, sample handling and analytics are outlined. Other important issues like the history of use, the hydrological connectivity between the treated area and the monitoring well, tracers, travel time estimations, groundwater flow direction, study duration and sampling frequency are addressed in different levels of detail. Because different ExAGs need different study designs (Section 2), it is not yet possible to develop harmonised guidance for regulatory purposes. The information provided in Gimsing et al. (2019) can, however, serve as a basis in developing detailed guidance for targeted monitoring studies.

6.1. Representative study designs

EC (2014) distinguishes between field leaching studies and groundwater monitoring studies. Gimsing et al. (2019) argue that the distinction between both study types is not always clear, particularly for in-field study designs. However, the paper does define the difference between field leaching studies and monitoring studies: 'Field leaching studies are usually conducted as a research study with carefully controlled agricultural operations including application of the active substance under supervision of the researcher, while monitoring studies are usually conducted in commercial fields where agricultural practices are controlled by the farmer'. The Panel agrees to this general definition and notes that field leaching studies can have in-field or edge-of-field study designs. The Panel notes that from a regulatory perspective, the distinction between both study types is highly relevant. Field leaching studies are carried out at Tier-3c and require in-context setting by inverse modelling (EC, 2014) whereas the endpoint of a groundwater monitoring study at Tier-4 is the measured concentration itself. The Panel notes that inverse modelling of field leaching studies is usually associated with a high level of uncertainty. The Panel therefore recommends reconsidering the role of inverse modelling in future guidance.

Gimsing et al. (2019) further distinguish between in-field and edge-of-field study designs for groundwater monitoring. In-field study designs generally benefit from the ease of proving connectivity, particularly for shallow monitoring depths. On the other hand, the Panel recognises that in-field study designs may be problematic due to concerns that artificial preferential flow pathways may develop along the well casing into the groundwater. Moreover, the risk of introducing topsoil containing residues to the gravel pack or the saturated zone remains, especially for newly installed wells. Both routes pose some risk of groundwater contamination due to the installation rather than leaching and may complicate the interpretation of the observations. The Panel notes that in-field study designs should be prospective rather than retrospective, because the latter provokes discussions whether peak concentrations were missed.

6.2. History of use

At all tiers, the groundwater risk assessment is based on the representative uses as defined in the GAP. Besides the timing of the treatment and the crop, the description of a use defines the maximum application rates and frequencies that are required for reasons of efficacy. As an active substance (or product) can be approved (or authorised) for several uses, the risk assessment needs to consider the pressures resulting from all intended uses in the various GAPs to meet the provisions of Article 31 and Annex II of Regulation (EC) 1107/2009. Therefore, the significance of groundwater monitoring results must explicitly be demonstrated for the application rates (in g a.i. per ha), timings (e.g. spring or winter application) and frequencies (e.g. biannual, annual, biennial, triennial application) that need to be assessed. Only if the link between the monitoring data and the uses according to the GAP is clearly established, the results can be considered a Tier-4 in the risk assessment.

The Panel notes that Gimsing et al. (2019) do not provide detailed recommendations how use data should be reported and assessed to demonstrate their compliance with the GAP. Several Member States comments on Gimsing et al. (2019) identified a lack of guidance in this regard. Though it was admitted that under practical agricultural conditions the strict boundary conditions as defined in a representative use as presented in the GAP table are seldomly fulfilled in a 1:1 manner in a groundwater monitoring study, they called for guidance to decide upon the question which application rates and frequencies the monitoring results are representative for. They expressed the need for clear decision criteria to conclude (i) if applications at a specific monitoring site adequately cover the intended use rate and frequency and (ii) how to proceed if they do not.

To obtain the historical use data might be a limiting factor in groundwater monitoring studies. Dependent on the travel times of the respective substances at the monitoring site, application data might be needed for a time span of many years (5–15 years). Long time series of data are especially required for strongly sorbing substances that move slowly through the soil column and the aquifer, either due to substance-specific (e.g. moderate or high adsorption to soil particles) or soil/aquifer-specific properties (e.g. high silt or clay contents). However, if product applications are not adequately reported, the measured concentrations cannot be correctly interpreted in relation to the GAP.

Consequently, a prerequisite for the regulatory use of targeted monitoring data is the availability of complete data on the use history of the products containing the respective active substances. The application data must be reported in sufficient detail for the whole monitoring period and the relevant preceding time span (considering the travel times of the substance). Note that generally time-series need to be longer for larger evaluation depths. This documentation needs to consider all fields with a hydrological connection to the monitoring wells. Reporting of sales and cropping data may in most cases not be sufficient, because these data can hardly be linked to a specific use on a specific field according to the GAP. The Panel notes that in addition to the targeted monitoring studies performed at Tier-4, results from public monitoring sites could be used to identify risk areas (see further Section 7).

Application rate

The application rate is an essential part of the definition of a use according to the GAP and thus in the groundwater risk assessment. In monitoring studies performed under agricultural conditions, the actual dose rates used might differ from the maximum rates in the GAP. Therefore, it is essential to report the application rates at every monitoring site in complete detail. Only if this information is available, an evaluation can be performed to designate which dose rates are covered by the monitoring results for risk assessment.

In recent evaluations in the context of active substance approvals at EU level, two different approaches were taken by RMS to assess targeted monitoring results based on different field application rates; (i) calculating a dose rate factor and (ii) up- and down-scaling of the results:

- In the RAR for the active substance S-metolachlor (June 2021), the RMS Germany calculated an average dose rate factor for each monitoring site for the period of the study. The average dose rate (considering only the years in which an application was done) was compared to the proposed application rates according to the GAP. For that purpose, a factor was calculated by dividing the average dose rate at each site by the intended dose rate. If this dose rate factor was 1 or above, it was concluded that the intended dose rate was covered at the respective site. If it was below 1, the site was not regarded as suitable to obtain information on the groundwater leaching potential for the respective dose rate. The evaluation of the monitoring results was then performed with sub-data sets for each representative use separately. This rather rigid approach has some limitations. The data set for the lower dose rates also includes the sites that were treated with higher dose rates and thus there might be a bias in the results. And if the actual application rates show high variations, the method can significantly reduce the number of acceptable sites covering an intended use.
- In the Addendum to the DAR for the active substance pinoxaden (April 2022), the RMS Austria introduced an approach of scaling the monitoring results with a rate-normalisation factor. This site-specific rate normalisation factor was calculated by dividing the intended dose rate by the actual averaged dose rate at the respective site. The factor was used for linear up- and down-scaling of the measured concentrations at each of the monitoring sites. For samples below LOD, the RMS suggested scaling based on ½ LOD.

The scaling method has the advantage that it is applicable to all monitoring sites and therefore maintains the number of sites of the whole data set. However, data-processing based on linear scaling ignores the impact of non-linear sorption processes and other non-linear phenomena occurring under field conditions. This can lead to significant error, particularly for substances with a Freundlich exponent < 0.9 (Table 4). The Panel therefore proposes to perform the scaling based on simulations with PEARL and/or PELMO for both the actual use, and the intended use. Simulations should preferably be done using soil and weather data for the respective monitoring site (this data is potentially available for the time-of-flight simulations described in Section 6.2). If insufficient local soil and weather data are available, a FOCUS scenario that matches the conditions at the monitoring site could be taken alternatively. The application frequency in the GAP should be used in the modelling (see below). The Panel acknowledges that these simulations increase the workload for both applicants, and regulators. Further guidance, including an operational procedure, must be developed in future guidance.

Application frequencies

Also, the application frequencies are part of the definition of a use according to the GAP and thus are an integral part of each tier in the groundwater risk assessment. Application frequencies specify both the number of applications per year and the minimum interval between applications and may further specify that applications in consecutive years are precluded. In simulating PEC_{GW} values at the lower tiers, different application frequencies can easily be considered. In FOCUS (2000), the impact of different application frequencies on the concentration in groundwater has been analysed. For two dummy substances, the 80th percentile concentration at 1-m depth was compared for annual, biennial and triennial applications modelled with PRZM, PELMO, PEARL and MACRO. In comparison to annual applications, the values were reduced by a factor of 2–3 for biennial applications and by a factor of 3–5 for triennial applications. The results of this modelling exercise indicate that less frequent application frequencies result in lower concentrations in groundwater.

Table 4: The 80th percentile of the annual average leaching concentration at 1-m depth (PEC) for the FOCUS Hamburg scenario and for FOCUS example substance D. The simulations were done with FOCUS PEARL 4.4.4

Results for FOCUS substance D with a Freundlich exponent (1/n) of 0.9			
Dose	**PEC with linear scaling[(a)]**	**PEC simulated with PEARL**	**PEC scaled divided by PEC simulated**
1.0 kg/ha	2.354 μg/L	2.354 μg/L	1.0
0.5 kg/ha	1.177 μg/L	0.879 μg/L	1.3
0.1 kg/ha	0.235 μg/L	0.095 μg/L	2.5
Results for FOCUS substance D with a Freundlich exponent (1/n) of 0.8			
Dose	**PEC with linear scaling**	**PEC simulated with PEARL**	**PEC scaled divided by PEC simulated**
1.0 kg/ha	0.481 μg/L	0.481 μg/L	1.0
0.5 kg/ha	0.241 μg/L	0.069 μg/L	3.5
0.1 kg/ha	0.048 μg/L	0.001 μg/L	48

(a): A simulation was done with an annual dosage of 1 kg/ha.

In legacy groundwater monitoring studies, data on product applications are often not reported in the level of detail required for regulatory purposes. To merely reference 'realistic agricultural conditions' covered by the actual application frequencies at Tier-4 is not sufficient. This would not be in line with the tiered approach that exactly considers the GAP at the lower tiers and would possibly underestimate the concentrations in groundwater under fields that receive more frequent applications of the respective pesticide.

In designing and conducting experimental leaching experiments (e.g. lysimeter or field leaching studies), applications can be (and have been) performed according to the regulatory needs (i.e. the uses as intended by the applicants). In groundwater monitoring studies, actual application frequencies in the treated areas at monitoring sites can be highly variable and it might be challenging to assign enough monitoring sites to the respective intended application frequency (or frequencies).

Due to the heterogeneity of the application frequencies at different sites in monitoring studies, conceptual approaches to link them to the use definitions in the GAP are not straight-forward. In the RAR for active substance S-metolachlor (June 2021), the RMS Germany split the entire set of monitoring sites into subsets considered more in line with, e.g. annual, biennial or triennial applications, e.g. as proposed in the GAP for S-metolachlor. However, this approach led to a distinct reduction of monitoring sites. Given the limited number of sites of monitoring studies and the different levels of assessment (EU, regulatory zones, Member States), a further reduction of the available data set should be avoided. The Panel therefore recommends developing guidance on how to relate the actual application frequencies observed in monitoring studies to the definition of the uses according to the GAP.

6.3. Connectivity

To demonstrate the hydraulic connectivity between the monitoring wells and the treated fields is challenging but crucial for a correct interpretation of the results, particularly if risk managers decide to set the evaluation depth at deeper values. It is needed to exclude false negative measurements, i.e. non-detections in the samples that are caused by the lack of product applications at the respective field. A proof of connectivity is also essential to demonstrate that the results obtained at a specific well have a significance for the field(s) for which the use history has been documented.

Gimsing et al. (2019) point out that for in-field studies with samples from the upper 10 cm of the groundwater, connectivity is essentially assured. Another approach to reduce the uncertainties associated with connectivity, that has been taken in several recent monitoring studies performed by applicants, is to install wells at the edge of the adjacent fields and to collect samples in the uppermost shallow groundwater (1- to 2-m filter screen installed near the groundwater table). This might be a suitable way to safeguard the sampling of maximum residues in groundwater originating from the targeted fields. An additional advantage of edge-of-field monitoring is that the risk of contamination of the samples is reduced (Section 6.1).

Establishing connectivity at larger spatial scales (e.g. the catchment of the aquifer) might be intricate and the data generated will seldomly be fit for the purpose to be used as Tier-4.

Tracers and pseudo-tracers

Control measurements are a long-established standard in many fields of experimental sciences to avoid misinterpretations of the results, e.g. to exclude false-positive and false-negative results. Though the practical restrictions in groundwater monitoring exercises are dissimilar compared to laboratory tests, the established principles of good science need to be followed. Against this background, the use of pseudo-tracers (e.g. metabolites) or tracers (e.g. bromide) is regarded as a suitable method to demonstrate connectivity between the field(s) that received the application (and for which use data is documented) and the sampling well.

This is in line with the suggestions of Gimsing et al. (2019) who summarised that tracers enable 'to determine flow connections/pathways, flow velocities and travel times, hydrodynamic dispersion, recharge, and discharge. [...] Findings of substances related to the target (e.g., primary or secondary metabolites) can act as 'tracers' providing proof of connectivity to soil treated with the target. In prospective studies, a conservative tracer such as bromide may be applied together with the target substance to provide proof of connectivity'. Several member state comments also support the use of tracers or pseudo-tracers as a proof of connectivity and demand that tracer experiments should become more important for targeted monitoring studies in the future to assist the hydraulic connectivity assessment.

Utilising metabolites as pseudo-tracers is, if possible, the simplest and most pragmatic way to prove connectivity. However, it must be considered that the presence of metabolites does not per se provide evidence that the monitoring well is only connected to one adjacent field in the delineated upstream area over time. The possibility that the well and its filter screen are temporarily also connected to another field that received any (unknown) kind of application with the active substance must be acknowledged in interpreting the results.

Human-applied tracers have pros and cons. The advantage is that they are applied at a known field. Thus, detections of the tracer in the monitoring well can unequivocally be assigned to originate from the respective treated field (this is, as described above, in contrast to pseudo-tracers). In monitoring studies with active substances, tracers might be the only tool available that guarantees appropriate reliability in demonstrating connectivity at a given site. On the other hand, there are several practical and formal obstacles that may impede the use of tracers. Gimsing et al. (2019) emphasised that such tracers 'should not be present in relevant concentrations in the hydrological system before the tracer experiment, retarded caused by sorption to or degradation in the soils/rocks, sensitive to changes in solution chemistry, toxic for the studied environment'. In most countries, legal permissions are required to treat whole fields with tracer substances like bromide; it might be challenging to obtain these permissions for a high number of sites located in different countries. This concern was also raised in the comment of CropLife Europe, indicating that applications of tracers to all fields may not possible due to: (i) permits needed, (ii) damage to crops and soil at commercial farms, (iii) applying at the same time as crop protection products may not be possible due to practical considerations and (iv) high background levels.

It must be kept in mind that any substance used as a tracer will behave differently compared to the monitored substance due to its specific environmental properties (e.g. degradation, adsorption) and consequently its travel time through the aquifer might be affected. Therefore, the tracer results need to be interpreted cautiously, especially regarding the temporal aspects of substance transport. This is particularly the case if the tracer is mobile compared to the substance under consideration. The Panel therefore suggests using time-of-flight modelling in addition to tracer and pseudo-tracer experiments.

Travel time (time-of-flight modelling)

A part of demonstrating connectivity to correctly interpret the measured concentrations is the temporal scale. Information about the travel time of the substance from the field through the unsaturated soil zone and the saturated groundwater zone to reach the sampling point must be available to establish sampling schedules that adequately represent the main contamination plume under the treated area. Estimation of travel times thus enables to predict if the measured residues can be interpreted as results of previous applications during the study period. In addition to the use of tracers, Gimsing et al. (2019) discussed these modelling approaches (so called time-of-flight simulations) to prove connectivity.

In the public consultation, the use of this method stimulated several comments. Member States stated that time-of-flight modelling provides a rough estimate in the absence of any other information. It provides a general idea of likely transit times under conditions of chromatographic flow, and to understand which sampling strategies might be appropriate. However, concerns were raised that it requires intensive data input and complex site-specific modelling that adds undesirable complexity for assessments of targeted monitoring studies. One comment took a critical position towards time-of-flight modelling and expressed that – in the light of the key position of proving adequate hydraulic connectivity in the overall evaluation of a monitoring study – neither a 'weight-of-evidence' approach, nor 'high-tech' modelling would be appropriate in a regulatory framework. The Panel agrees that results from time-of-flight analyses are subject to considerable uncertainty, but also considers time-of-flight modelling an essential part of the analysis of hydraulic connectivity. Because of the uncertainty associated with time-of-flight modelling, the Panel suggest to additionally perform tracer and/or pseudo-tracer experiments, as mentioned above. The combined use of tracer experiments and modelling can in the view of the Panel provide sufficient evidence to demonstrate hydraulic connectivity. The uncertainties of the modelling approaches are discussed in the remainder of this section.

Time-of-flight simulations are a combination of one-dimensional pesticide leaching models and three-dimensional groundwater transport models. Results of time-of-flight simulations depend on the assumptions made and on the quality of the input data. The one-dimensional models (e.g. PEARL, PELMO and MACRO) simulate leaching in the unsaturated soil zone, PEARL and MACRO also simulate transport in the uppermost saturated zone. These models have options to simulate preferential flow; however, in the lower tiers of FOCUS groundwater chromatographic flow is assumed (except for one scenario). When used for time-of-flight analysis, the arrival time of a substance is a critical parameter. Chromatographic flow models may not be suitable for predicting early arrival times of pesticides in situations where preferential flow is an important transport mechanism (this is acknowledged also by Gimsing et al., 2019). Some authors have used larger dispersion coefficients to account for preferential flow; however, this is not recommended because unrealistic dispersion coefficients would be needed to simulate early arrival times of pesticides (e.g. Scorza-Júnior and Boesten, 2005).

Another critical aspect of site-specific simulations regards the fate properties of the substance such as the degradation half-life and adsorption parameters. In the modelling 'average' values are used, although it is known that these parameters are location dependent (Section 4.3) and show considerable variability (Walker and Thompson, 1977; Allen and Walker, 1987). To correctly simulate the travel time in the unsaturated zone at a given site, it would be required to determine site-specific sorption parameters. Kupfersberger et al. (2017) studied the impact of this phenomenon and concluded that if no physicochemical analysis of local soils is available, the use of 'average' values for environmental fate parameters might lead to only a rough description of the leaching features for a given compound or unrealistic fitted parameters that will not be appropriate for predictions. To deal with this uncertainty, the Panel recommends calculating the travel time using both the lowest and the highest sorption coefficient to determine the range of possible travel times. The Panel recommends developing guidance for adequate travel time estimations, considering the uncertainty of the sorption coefficient. The procedure in EFSA PPR Panel (2012) could serve as an example. Notice that uncertainty of substance properties is also considered in the in-context setting procedure (Section 4). However, because we are interested in the range of possible travel times, a different solution to handle this uncertainty is chosen here.

Dispersion is another process that affects the width of the solute breakthrough at any depth, with broader breakthrough curves for larger values. It is well known that dispersivity is scale-dependent, with increasing values at larger scales (Vanderborght and Vereecken, 2007). With that in mind, an extension of the soil profile in the time-of-flight analysis to larger depths goes along with an increase of the value of the dispersivity. However, the implementation of the values for the dispersivity in time-of-flight simulations presented in appendix 6 of Gimsing et al. (2019) is premature. A justification for the absolute values of the enlargement of dispersity over depth is required based on a literature review or on simulations. Therefore, the Panel recommends that the implementation of a larger dispersivity for a time-of-flight analysis needs to be further analysed before the implementation in guidance.

Commonly, one-dimensional leaching models like PEARL and PELMO are utilised to estimate leaching concentrations at a depth of 1 m. In time-of-flight simulations, predictions for deeper soil layers might be needed to adequately represent the conditions at a given site. This extrapolation to depths below 1 m requires accuracy, as it implicitly affects the prediction of the travel time through the soil column. However, parameterisation of the models for simulations of deeper soil layers is challenging as the required data for depths below 1 m might not be available for all sites. And even if these data are collected during the study phase, this might be too late to utilise it in elaborating adequate sampling schedules.

To estimate the horizontal flow and transport within the saturated zone, three-dimensional groundwater transport models can be used. Gimsing et al. (2019) discussed the assumptions and parameterisations of the available models (e.g. Feflow, Modflow and OpenGeosys) and questioned the current formal suitability of these models, as they are not mentioned in the existing guidance documents. As for modelling vertical flow with the one-dimensional models in deeper layers than 1 m, the availability of the input data required for the simulations might be challenging.

In the regulatory authorities in the Member States, the practical expertise using these software tools is limited. However, for a transparent and critical assessment of time-of-flight estimations, the Member State experts must be enabled to reproduce the evaluations submitted by the applicants. Therefore, these tools must be publicly available, user-friendly and under version-control (e.g. FOCUS version control). The Panel recommends following guidance in the Good Modelling Practice Opinion (EFSA PPR Panel, 2014).

One of the indicators in time-of-flight modelling is the months at target concentration (MTC), which is a measure to ensure that concentrations within a certain percentage of the peak concentration will be observed. The Panel recommends a fundamental discussion for future guidance, whether the proclaimed 70% of the peak concentration is a good measure for the minimum time window where the substance is above the defined target concentration. In terms of reporting, all listed measures in appendix 6 of Gimsing et al. (2019) should be documented for the single years and an aggregation of the overall results in terms of the minimum, average and maximum MTC should be provided.

Groundwater flow direction

A stable groundwater flow direction over time ensures that the monitoring well potentially covers groundwater residues from the adjacent field. The groundwater flow direction can change over time. This can be the case especially in areas with very shallow groundwater, which are – on the other hand – preferable sites as here the travel times through the unsaturated zone might be short enough for the duration of a monitoring study.

Gimsing et al. (2019) emphasise that '[...] the direction of groundwater flow is needed to optimally locate the monitoring wells. Therefore, unless the groundwater flow direction is obvious from the slope of the land and the position of water bodies, three wells will typically be installed at the start of the study to determine (or confirm) the direction of groundwater flow. Groundwater flow may need to be checked regularly during the study since the flow may change direction with time in some locations, so additional wells may be installed if required. Usually, after the groundwater flow direction is determined 1–10 wells down gradient and in some situations, wells may also be installed upgradient of the treated field to determine if an active substance or its metabolites are present in groundwater flowing into the field from adjacent fields'.

Various comments dealt with the uncertainty related to fluctuations in the groundwater flow direction. Member States highlighted that due to variable groundwater flow directions, nominated sampling wells may be temporarily or even permanently unconnected to a treated field, particularly at monitoring sites where multiple sampling wells have been installed to address variable groundwater flows directions. It was pointed out that precise information on the groundwater flow direction over time and the hydrological conditions in the unsaturated and the saturated zone at a chosen site is usually limited at the beginning of the monitoring period and that constant measurements of the groundwater flow direction should become more important for targeted monitoring studies in the future. Clear criteria were requested on when to consider a sampling well adequately connected and when not, and how to proceed if a sampling well is not adequately connected (e.g. by accounting for a minimum percentage of samples required with the sampling well adequately located in the groundwater flow direction based on groundwater contour plots obtained at each sampling date).

At locations with an alteration in groundwater flow direction, the delineation of the catchment field area corresponding to the monitoring wells is called into question. If the direction of the groundwater movement changes during the study phase, a different field may constitute the 'catchment' area. It may therefore be difficult to relate the measured concentration to the actual product use, which is essential when monitoring studies are used at Tier-4. Accordingly, even if connectivity can be proven by means of the occurrence of metabolites or tracers, uncertainty remains if the maximum concentrations of a compound in the residue plume is adequately captured over time.

Due to these considerations, the Panel recommends recording groundwater levels at enough wells/piezometers at each monitoring site (at least 3) in a high temporal resolution (e.g. daily) to allow the observation of the groundwater flow direction over the whole period of the study. Guidance must be developed to decide how to identify relevant fluctuations in the groundwater flow direction at a given site and how to consider them in the evaluation of the monitoring data.

6.4. Sampling frequency

Gimsing et al. (2019) do not provide specific recommendations on sampling frequencies. In case of limited knowledge about the hydrogeological situation in the unsaturated and the saturated zones, monthly sampling is recommended; in some of the studies provided as illustrative examples in the paper sampling was performed quarterly. Member States acknowledged in their comments that the sampling frequency depends on several factors like mobility and persistence of the active substance or metabolite, hydrological site characteristics, climate conditions, depth to groundwater, filter screen depth, application time, prospective and retrospective study design and the ExAG. It was also remarked that even when there is a good knowledge about the hydrogeological regime of the saturated zone, a sampling effort (monthly) should be done at least for the first years of a monitoring study to better understand the temporal variations. Sampling interval can then be adapted depending on results from these first years. The Panel notes that guidance is needed on temporal aspects of sampling strategies.

In fact, the issue of sampling frequencies for monitoring studies touches interconnected questions of setting ExAGs and practical aspects of study design: The frequency of sampling determines if temporal or seasonal effects are detected and thus the temporal resolution of the sampling schedules is directly related to the temporal aspect of the ExAG, i.e. if seldom, regular or frequent trigger value exceedances are considered relevant.

On the other hand, there are numerous substance- and site-specific parameters that must be considered in deriving a sampling schedule. Often, the sampling frequency is set at the beginning of the study while the information required to scientifically derive a site-specific sampling schedule is often collected during the study phase, which might be too late.

The temporal dimension of the relevant concentration, the multiyear temporal statistical population of relevant concentration values for one spatial unit, and spatiotemporal percentile of the statistical population of the relevant concentration are elements of the ExAG for protection of groundwater (Table 2). Temporal aggregations of the concentration observations are required per monitoring site in terms of descriptive statistics (e.g. mean or percentile values) that comply with the ExAG to be defined. The temporal aggregation requires statistical techniques that are based on survival or reliability analysis methods when left-censored observations occur in the data set (see section on left-censored data in Section 8).

6.5. Number of monitoring sites

For the EU approval of active substances at EU level, a 'safe use' for at least one of the climatic zones represented by the nine FOCUS groundwater scenarios must be demonstrated (EC, 2014). Typically, the product authorisations in the Member States are not only based on the results of one or several FOCUS scenarios, but also consider the specific agricultural and environmental circumstances. Currently, harmonised approaches to translate the one safe-use concept of the lower tiers to Tier-4 are not available.

The FOCUS scenarios aim to be at the 90th percentile vulnerability (EC, 2014). The FOCUS group proposed that 90% of analyses, obtained from at least 50 locations would need to be below the regulatory threshold for the Commission to consider a proposal for EU level approval (a location is defined as a single well or group of wells at the same site). A smaller number of locations (approximately 20) would be acceptable if they are specifically targeted to the pesticide of interest. However, the FOCUS group admitted that there was no statistical basis for these number of locations. In the peer review for active substance S-metolachlor (EFSA, 2023), a minimum number of 20 sites for each FOCUS climate zone was considered to obtain representative data for each of the agricultural regions of the nine FOCUS zones. Also, this decision was not based on statistical evaluations.

In the development of the bee guidance document, a simulation exercise was conducted to define the minimum number of locations for field exposure studies (Revised draft guidance on the risk assessment of PPPs on bees, Annex B, 2022). The analysis is based upon a log-normal distribution of the main driver of the target variable (i.e. the daily residue intake per bee). The Panel recommends reviewing if a comparable methodology can be developed to statistically determine the minimum number of sites for groundwater monitoring studies. This analysis needs to consider the specific requirements of this study type regarding substance- and site-specific leaching characteristics. It should also consider the combination with suggested vulnerability mapping procedures, because this can reduce the number of required groundwater monitoring sites (Section 4). Finally, the definition of the minimum number of sites needs to address the suitability of the results for the different levels of assessment (EU, zones, Member States).

The statistical population of spatial units and spatiotemporal percentile of the statistical population of the relevant concentration are elements of the ExAG for protection of groundwater (Table 2). Spatial aggregations of temporally aggregated concentration observations are required in terms of descriptive statistics (e.g. mean or percentile values) that comply with the ExAG to be defined. The spatial aggregation requires techniques that are based on survival or reliability analysis methods when left-censored observations occur in the data set (see section on left-censored data in Section 8).

6.6. Fixed monitoring networks

In their comments, a few Member States expressed their concern about the complexity of targeted monitoring studies. AGES argued that 'a harmonised study design is a key for regulatory acceptance and provides the basis for a harmonised evaluation. In this respect, AGES recommends a rather rigid study design for targeted edge-of-field monitoring studies with newly drilled monitoring wells, giving clear advice on, e.g., the minimum/maximum field size, minimum number of sampling wells, minimum number of supporting wells (assisting groundwater flow direction assessments), minimum depth to groundwater, minimum sampling frequency, minimum study duration, etc.'. The Panel agrees that a 'rigid study design' helps regulatory acceptance of groundwater monitoring studies. A possible way forward would be establishing a network of standardised monitoring sites. The Panel therefore recommends investigating if such a monitoring network could be established at EU level.

The Danish Pesticide Leaching Assessment Programme (PLAP)[20] could serve as an example of a fixed monitoring network. Currently, this programme encompasses six fields that represent dominant soil types and climatic conditions in Denmark (Rosenbom et al., 2015). The groundwater table is relatively shallow at all the fields, enabling rapid detection of pesticide leaching to the groundwater. Cultivation of the PLAP fields is done in accordance with the conventional agricultural practice in the local area. The pesticides are applied at maximum permitted doses as specified in the GAP tables. Thus, any pesticides or degradation products appearing in the groundwater downstream of the fields can, with a few exceptions, be related to the current approval conditions and use of the given pesticide. Since the start of the programme, 52 pesticides and 99 degradation products have been included in PLAP. These pesticides have been selected based on expert judgement by the Danish EPA (Badawi et al., 2022). Notice that PLAP fields are basically field leaching studies, although they follow an edge-of-field study design.

Fixed monitoring programmes have several advantages. First, the set-up of a monitoring network only needs to be done once, increasing efficiency of the regulatory process. Second, the study sites can be selected according to commonly agreed principles using the latest techniques (computation time of models used for site selection is not an issue and the work can be done by independent scientists). The Panel considers selection of representative and vulnerable sites an important condition for any monitoring study to overrule the early tiers of FOCUS groundwater. Third, hydrological and site conditions, as well as management practices and pesticide use are well known. The implementation of

20 http://pesticidvarsling.dk/?lang=en

a network of fixed monitoring sites would require extensive efforts in the planning and installation phase but could potentially prove to be cost-effective and resource-saving in the long term (as e.g. the site selection and monitoring well installation would need to be performed only once and not for every single study).

These types of study also have disadvantages, the most important one being that only a few sites can be managed. Furthermore, unexpected hotspots of leaching could be missed when monitoring a few sites only. A possible way to mitigate this problem would be to set up a harmonised public groundwater monitoring system in addition to the fixed monitoring programme (see Section 7 for requirements of such a monitoring programme). Another disadvantage of fixed monitoring sites is that they may not be sufficiently conservative for all substances. To mitigate this problem, the Panel proposes selecting the monitoring sites based on a wide range of substances. A similar approach has been developed for the selection of the exposure scenarios for soil organisms (EFSA PPR Panel, 2012; Tiktak et al., 2013). Nevertheless, in-context setting by an appropriate vulnerability analysis may still be necessary, particularly for substances whose properties depend on other soil properties than organic matter (e.g. pH).

6.7. Conclusions and recommendations

- A prerequisite for the regulatory use of monitoring data is the availability of complete data on the use history of the products containing the respective active substances. The application data must be reported in sufficient detail for the whole monitoring period and the relevant preceding time span (considering the travel times of the substance).
- The Panel recommends developing guidance how to relate the actual application frequencies observed in monitoring studies to the definition of the uses according to the GAP. In this respect, the so-called scaling method has the advantage that it is applicable to all monitoring sites and therefore maintains the number of sites of the whole data set. The Panel recommends scaling based on simulations with PEARL and/or PELMO for both the actual use, and the intended use.
- The combined use of measurements of pseudo-tracers (e.g. metabolites) or tracers (e.g. bromide) with site-specific time-of-flight modelling exercises is regarded as a suitable and practical approach to demonstrate connectivity between the field(s) that received the application (and for which data on use is documented) and the sampling well.
- Results of pseudo-tracers and tracers must be interpreted with care, e.g. if the substance under consideration has a higher sorption coefficient than the tracer.
- Time-of-flight simulations are associated with a high degree of uncertainty due to for example a lack of input data for deeper soil layers. Variability of pesticide properties adds additional uncertainty to the prediction of travel times. To deal with this uncertainty, the Panel recommends calculating the possible range of travel times at a site, considering variability of sorption coefficients.
- Time-of-flight modelling requires a high degree of experience with hydrological modelling and site-specific model parameterisation. Further guidance and training courses would be needed if these approaches become more important in regulatory use.
- Information about the flow direction is one of the major requirements of a monitoring site to exclude false-negative measurements in the interpretation of the monitoring results. It is a prerequisite for assessing the suitability of a site for leaching assessments and thus determines its reliability in the regulatory risk assessment. Guidance needs to be developed to decide how to identify relevant fluctuations in the groundwater flow direction at a given site and how to consider them in the evaluation of the monitoring data.
- The Panel recommends developing a statistical approach to determine the minimum number of sites required for regulatory groundwater monitoring that considers all levels of assessment (EU, regulatory zones and Member States).
- A suitable sampling frequency depends on different site-specific conditions (soil, weather, hydrology, product use) as well as substance specific properties. The Panel recommends considering this aspect when guidance is developed for regulatory use.
- A network of standardised monitoring sites could help regulatory acceptance of groundwater monitoring studies. The Panel recommends investigating if such a monitoring network could be established at EU level.

7. Requirements for using public monitoring programmes

Monitoring data collected by third party organisations, such as environmental agencies, water companies and research institutions, are designed for purposes other than for the authorisation of PPPs under Regulation (EC) 1107/2009. The quantity and quality of public monitoring data on active substances and their transformation products can vary strongly and may thus limit their potential use in the regulatory risk assessment. However, the data set of the substance of interest, possibly merged from different monitoring programmes, cover larger temporal and spatial scales than targeted monitoring studies. This means that concentration data in groundwater bodies is available for many aquifers under a wide variety of environmental conditions and agricultural practices. Results from public monitoring might unveil substances of concern and/or specific areas of high vulnerability. They can also illustrate trends over time. Thus, flaws of publicly available monitoring data, e.g. the likely impossibility to identify and exclude false negatives (i.e. the active substance was not applied in the upgradient area of the well), is weakened by a vast amount of available monitoring data. This is a great advantage of public monitoring programmes.

7.1. Quality criteria

At present, publicly available monitoring data often do not meet the main quality criteria for a Tier-4 risk assessment, e.g. the reported use pattern of plant protection products in the upgradient area of the wells, the hydraulic connectivity and a robust estimate of solute travel times between the soils in the study area to the sampled aquifer. The Panel recommends not to use publicly available monitoring data as a Tier-4 for authorisation of PPPs to demonstrate safe use, i.e. for overruling the lower tiers. Nevertheless, publicly available monitoring data provide important information on the current state of and possible trends in groundwater bodies and provide additional information on the applicability and interpretation of targeted monitoring studies. Therefore, the Panel recommends including publicly available monitoring data in the risk assessment as supportive information. If the monitoring data shows, for example, large-scale exceedances of the drinking water limit for a substance that has passed the lower tiers of the groundwater leaching assessment, applicants could be asked to provide additional information to demonstrate the safe use of the product. To improve access to the monitoring data, harmonised databases could be helpful. For this purpose, the Netherlands developed the so-called 'Groundwater Atlas'. The Atlas not only provides access to the data, but also to harmonised metadata and quality criteria. A methodology for the use of these monitoring data in the leaching assessment has been proposed together with the publication of the Groundwater Atlas. The aim of the procedure is to provide a plausible relationship between the presence of a substance in groundwater and the authorised use (Kruijne et al., 2023). The aim of the scientific paper by Gimsing et al. (2019) was to establish scientific recommendations for conducting and interpreting groundwater monitoring studies with no intention to be a guidance document. In their comments, Member States raised issues concerning the assessment of publicly available monitoring data that need to be additionally addressed in a guidance document.

To include publicly available monitoring data as supportive information in the risk assessment, the data used need to comply with minimum quality criteria on site characterisation, sampling strategy, data interpretation and reporting. The listed quality criteria for publicly available monitoring data in Gimsing et al. (2019) is neither exhaustive nor conclusive. A proposal for a list of items on how publicly available groundwater monitoring data should be collected and characterised is provided by the European Groundwater Monitoring Network for Regulators and endorsed by the Panel (see Appendix C). This list of items is an extension of the list provided in Chapter 7.3 of Gimsing et al. (2019) and can be used as a basis on how to report the relevant information needed to characterise public monitoring data for active substance approvals under Regulation (EC) 1107/2009.

7.2. Shared transformation products

Risk assessment of active substances under Regulation (EC) 1107/2009 follows a 'one substance-one assessment' approach. However, active substances may share the same transformation product(s). Shared transformation products may arise from legacy or registered plant protection products. Examples are (i) triazine amine, which is a common metabolite of various active substances of the sulfonylurea group, including prosulfuron, metsulfuron-methyl, thifensulfuron-methyl, triasulfuron, and iodosulfuron- methyl-sodium; (ii) two common metabolites of active substances belonging to the group of succinate-dehydrogenase inhibitors, including bixafen, sedaxane, fluxapyroxad, benzovindiflupyr and isopyrazam; (iii) 1,2,4-triazole from conazole active substances and (iv) common pyrethroid metabolites (EFSA PPR Panel, 2022).

A transformation product of active substances may also be identical with a chemical having another source of origin, with trifluoroacetic acid (TFA), 1,2,4-triazole and aminomethylphosphonic acid (AMPA) being examples. TFA may be a common metabolite of a range of approved active substances, such as fluazinam and flufenacet, but can also originate from other sources where TFA is used in synthesis processes as a starting material or as a solvent at an industrial scale. AMPA, a metabolite of glyphosate has been postulated to be formed from phosphonate detergents in wastewater treatment plants so might enter agricultural fields when sewage sludge is used as agricultural fertiliser.

During the harmonisation of the EFSA endpoints, studies already reviewed in the EU with the shared metabolite are used to derive the mean of the parameter values in lower tiers of groundwater risk assessment. At Tier-4, there is no recommendation in Gimsing et al. (2019) to interpret the monitoring data of shared metabolites. The Panel recommends addressing this issue in a future guidance document.

7.3. Conclusions and recommendations

- Publicly available monitoring data provide important information on the current state of and possible trends in groundwater bodies and provide additional information on the applicability and interpretation of targeted monitoring studies due to vast amount of available data. Therefore, the Panel recommends including publicly available monitoring data in the risk assessment as supportive information.
- At present, publicly available monitoring data often do not meet the main quality criteria for a Tier-4 risk assessment. The Panel recommends using these data as supportive information and not to use these data to overrule earlier tiers but to possibly trigger the need for further information by the applicants.
- To include publicly available monitoring data as supportive information in the risk assessment, the data need to comply with quality criteria. These quality criteria must be established in a future guidance document.
- Guidance is needed on the interpretation of monitoring data of shared metabolites.
- Access to the data for regulators could be improved by making them available in a standardised European database, together with the metadata and procedures to establish a plausible relationship between the presence of the substance in groundwater and the authorised use.

8. Reporting results from groundwater monitoring studies

Chapter 6 in Gimsing et al. (2019) give some considerations for the study report of monitoring studies to be submitted to regulatory authorities. A list of items to be reported is included in section 5.2 of Gimsing et al. (2019). This list is, however, not comprehensive for targeted monitoring studies. The Panel suggests creating a list of items comparable to the list of items to characterise publicly available monitoring data (see Appendix C). The monitoring report should be structured as follows:

- A description of the set-up of the monitoring study, including results from the site selection and the in-context procedure,
- Monitoring results from all individual monitoring sites including a critical evaluation of results,
- A description of the aggregation procedure to derive the regulatory endpoint, including an evaluation of the result,
- A discussion section, addressing the most important uncertainties of the study,
- Conclusions.

The Panel agrees with Gimsing et al. (2019) that the ExAG affects the design of the study and the procedure for deriving the regulatory endpoint. Therefore, a final list of requirements can only be made when the ExAG has been defined by risk assessors after consultation of risk managers (Section 2). Gimsing et al. (2019) mention that 'the report must provide a sufficiently detailed description of materials and methods to understand what was done in the study and allow others to reproduce the experiment'. The Panel notes that criteria must be further elaborated when developing the guidance document.

8.1. Description of the set-up of the monitoring study

This section should first describe the objective/aim of the monitoring study, followed by the description of the site selection and the description of the actual design of the monitoring sites.

Objective/aim of the monitoring study

The description of the objective/aim of the monitoring study should include the following elements:

- The risks identified at the lower tiers of the groundwater leaching assessment and justification why a monitoring study needs to be performed;
- Properties of the active substance and metabolites;
- The intended product uses. This should include the crops in which the product is intended to be used, the maximum application rate of the product, the frequency and time of application, and the type of application (see also Section 6.2). Also, the geographical area for which an approval/authorisation is requested must be specified;
- The ExAG. Until a harmonised ExAG has been derived at EU level, it must be discussed with the regulatory authorities before setting up the monitoring programme. All six elements of the ExAG described in Section 2 of this Statement must be addressed.

Site selection

Site selection consists of two phases, i.e. selection of vulnerable regions (preselection phase) and the final selection of the monitoring sites. Both depend on an appropriate vulnerability analysis (Section 4). In the final selection of the monitoring sites, more detailed information is needed, e.g. a time-of-flight analysis (Section 6). The report must contain:

- A description of the model used for the vulnerability analysis. If FOCUS models are used, reference to the version number is sufficient. If non-FOCUS models are used, a justification for using this model must be submitted as well (Section 4.3);
- A description of the input data for modelling. This includes among other substances properties, management practices, soil data, climate data and crop data (see further Section 4.3);
- Results of the preselection phase. The identified regions should be in typical agricultural areas representative of the intended product use. The regions must further meet the agreed ExAGs, i.e. these regions should be sufficiently vulnerable to pesticide leaching (Section 4.4);
- Results of the final site selection. Properties of the selected sites must be representative of the regions identified in the preselection phase. If multiple substances are monitored, site selection can be based on the substance with the most important concentration exceedances (Section 4.4); however, a justification must be given in this section.
- Results from the in-context setting procedure must be reported, including a description of the models and input data used for the in-context setting procedure. Additional selection criteria based on the time-of-flight analyses and an inventory of product use in the catchment area need to be reported as well (Section 6). Sites that do not meet the selection criteria need to be discarded.

Design of monitoring sites

The report must include the following details related to design of monitoring sites:

- A description of the study design (in-field, edge-of-field), including the number of monitoring locations and wells per monitoring site, depth of the filter screens and exact location of the wells (Section 6). Study design must be in line with the ExAG and must be justified;
- A description of the monitoring sites covered by the monitoring well(s). This includes – for each site- the exact geographical location, management practices (including irrigation and drainage), information on product use, information on other activities that may affect the observed concentration, analysis of the groundwater direction, the soil profile, groundwater depth and aquifer type (see also the list in section 6.2.3 of Gimsing et al., 2019);
- Sampling strategy/procedure and analytics (see Appendix C for these items).

More details are given in Section 6 of this Statement.

8.2. Reporting of results from the individual monitoring sites

To get a full picture of the leaching risk from a monitoring study, all measurements must be presented. The following information must be provided (see also Appendix C):

- the total number of measured samples;
- the total number of wells sampled;
- the number and percentage of samples and wells above the limit of quantification (LOQ);
- the number and percentage of samples and wells above the regulatory threshold (usually 0.1 μg/L);
- for non-relevant metabolites, the number and percentage of samples and wells above 0.75 and 10 μg/L must be reported as well.

Note that the number and percentage of samples is not the same as the spatial and temporal aggregate in the regulatory endpoint, as defined in the ExAG (see further Section 8.3). If applicable, the observed concentrations must be scaled to the maximum dosage in the GAP and consider the difference in application frequencies, such as within the year or between years (see Section 6.2 of this Statement for details). Results from the scaling procedure must be presented and described. A discussion on uncertainties is to be submitted together with the data (Section 8.4).

In line with commonly used regulatory 'shorthand' terminology, Gimsing et al. (2019) made use of the terms 'false positives' and 'false negatives'. This same terminology is used in this statement for consistency with the paper being reviewed. However, it should be noted that reliably sampled groundwater, where best sample handling practice and validated analytical methods have been used for analysis, does not provide false results, but rather represents reality regarding each sample taken in the specific context. So, the context of the way these terms have been used is that outlined below, such that false relates to what sample results from a location need to inform about and not that any results are wrong.

False positives

The study report must discuss for each site if the measured concentrations result from the intended uses in the GAP. Unexpectedly high concentrations or irregular concentration patterns in groundwater samples often trigger in-depth examinations of the respective sites and their surroundings. The aim is to elucidate if there is justified evidence that processes other than leaching caused these unexpectedly high concentrations (so-called 'false positives'). Such processes can result from point entries caused by shortcuts into groundwater, due to e.g. farmyard runoff, damaged wells or well capping, construction works in the field area. Moreover, specific agricultural practices like tillage into deeper soil layers due to the removal of wooden roots in tree nurseries that might cause preferred pathways, or the influence of surface waters have been mentioned by applicants to justify the exclusion of sites from the data sets.

While Member States commented that the identification of false positives might be 'extremely challenging, particularly for evaluators having limited access to data', the Panel considers such elucidations useful to guarantee that results of the respective sites provide information about the leaching behaviour of the substance. If there is convincing evidence to unequivocally conclude that the groundwater contamination at a given site is the result of other processes than leaching by chromatographic flow or transport by preferential flow, the Panel considers exclusion of such sites acceptable. The Panel notes that groundwater monitoring is currently the only way to potentially detect the occurrence of preferential flow and therefore considers exclusion of sites with preferential flow unacceptable. Exclusion of monitoring sites must always be supported by verifiable supporting information.

False negatives

The detection of false negatives must also be discussed. False negatives can result – among others – from lack of hydraulic connectivity between the field and the monitoring well, incomplete or erroneous documentation of product applications, local travel times at the site deviating from predictions, dry weather conditions, dilution by non-agricultural upstream areas, clogged filters, storage and stability issues or analytical problems. Consequently, the concentrations measured at a site during the study phase might be permanently or temporarily low, whereas they could have been higher if one of the above-mentioned circumstances had not appeared. A critical evaluation of site characteristics is therefore needed for measurements below the limit of quantification. This aspect needs further attention when developing regulatory guidance.

8.3. Derivation of regulatory endpoints

For the derivation of regulatory endpoints, the observations in the single wells need to be aggregated in time and in space in an appropriate way. These aggregations should comply with the six elements of the ExAG for the protection of groundwater (see Table 2). Observations from single monitoring sites, i.e. a spatial unit, are summarised according to the specified temporal dimension of the relevant concentration and reported in terms of the number (or percentage) of concentration data that belong to a certain class. These classes represent number (or percentage) of observations, for example, below the limit of detection or quantification or below or above some defined numerical value (e.g. 0.1, 1.0, 2.5 or 10 μg/L). Then, the multiyear temporal statistical population of relevant concentration values for one spatial unit is derived. This is repeated for all spatial units within the population of interest. Finally, a spatiotemporal percentile of the relevant concentrations can be deduced, which is used as a regulatory endpoint.

Note that in the aggregation procedure, not all measurements have equal weight, particularly when they are unevenly distributed throughout the averaging period or do not sample the same depth. Let us assume, for example, that the temporal dimension of the ExAG is an annual average concentration. Let us further assume that we have one measurement in the period January–March, two measurements in the period April–June, three measurements in the period July–September and no measurements in the period October–December. In this case, taking a simple arithmetic mean of these six measurements would not represent the temporal dimension of the ExAG appropriately. This example shows that guidance is needed on how to aggregate individual measurements to a regulatory endpoint. This guidance should also consider how to deal with missing data.

Left-censored data

Descriptive statistics, such as the mean or percentiles, are of interest when groundwater monitoring data need to be aggregated temporally or spatially to comply with the ExAG to be defined. This raises the question how to deal with left-censored observations. Statisticians refer to values below or above a threshold as censored observations. Helsel (2012) defines left-censored observations in the field of environmental sciences as 'low-level concentrations of organic or inorganic chemicals with values known only to be somewhere between zero and the laboratory's detection/reporting limits'. Concentration data from groundwater monitoring studies are inherently left-censored data as the result of analytical limitations, with possible values below limit of detection (non-detects), between the limit of detection and quantification (not quantified), or above the limit of quantification (detects) in the data set. Left-censored observations must be considered in the analysis of descriptive statistics (mean, percentiles) because they contain information. Ignoring or fabricating censored observations, the latter being the substitution of the non-detects with some kind of constant value (e.g. half the limit of detection), introduce a bias and a non-existing pattern into the data set and no reliable, i.e. biased, estimates of the descriptive statistics are possible (Helsel, 2006).

Techniques that are based on survival or reliability analysis methods can deal with left-censored observations. Depending on the underlying shape of the concentration distribution, the number of data points and the percentage of non-detects, methods such as maximum likelihood estimation, Kaplan–Meier, or regression on order statistics, among others (Helsel, 2012; Shoari and Dubé, 2018), should be applied. This class of methods can also deal with the change of detection limits over time in a data set.

The Panel advocates the use of techniques that are based on survival or reliability analysis methods for the descriptive statistics to deal with left-censored observations and provide guidance on this issue.

8.4. Discussion and conclusions from the monitoring study

Uncertainty analysis

Annex II of Regulation (EC) No 1107/2009 states that the assessment should consider the 'uncertainty of the data'. This implies that explicit consideration of uncertainty is appropriate, both by the PPR-Panel and/or when developing a guidance document for groundwater monitoring studies, and during the assessment of specific risks (see Section 2 of this Statement).

According to EFSA PPR Panel (2010), the tiered system must address the risk assessment with greater accuracy and precision when going from lower to higher tiers. Consequently, the uncertainty analysis should consider the main sources of uncertainty of groundwater monitoring studies and highlight the dissimilar degrees of uncertainty at the different tiers of the groundwater risk assessment and their expected overall impact on the assessment outcome. The Panel considers the analysis of possible false negatives to be part of the uncertainty analysis (Section 8.2). Sources for uncertainty can be identified in all phases of a monitoring study, including:

- Vulnerability analysis for site selection and in-context setting. Uncertainties may arise both from the modelling approach, and the data used for modelling (Section 4);
- Availability of active substance application data and scaling to the application rate in the GAP (Section 6);
- Impact of alterations in the groundwater flow direction (Section 6);
- Connectivity assessment based on time-of-flight analyses or tracer/pseudo-tracer experiments (Section 6);
- Prediction of travel times and the depth of the filter screens relative to the depth of the groundwater table (Section 6);
- Installation of the monitoring wells (Section 6);
- Sampling (Section 6);
- Analytical methods and handling of samples (Section 6);
- Variability of weather conditions during the monitoring experiment (Section 6).

Notice that this list is not exhaustive and needs to be further developed in the final guidance document.

The methodology suggested in EFSA SC (2018) could form the basis to systematically review major sources of uncertainty. Sources of uncertainties must be addressed qualitatively, and where possible quantitatively (see e.g. Section 4 of this Statement where a procedure to quantitatively deal with uncertainties in pesticide properties is suggested). The Panel further suggests that guidance needs to be developed on how to use results from this uncertainty analysis in regulatory assessments. Of particular importance is the question which level of certainty would be needed to justify a 'safe use'. Note that the installation of monitoring wells, problems with sampling and variability of weather conditions can invoke 'false positives' and 'false negatives'. A balanced discussion, in which both these 'false positives' and 'false negatives' are addressed, should be part of the regulatory report (Section 8.2). So, the uncertainty analysis should indicate whether the different sources of uncertainty are likely to underestimate or overestimate the concentration in the field.

Conclusions of the study

The study report should include a conclusion whether a 'safe use' for the intended use and/or GAP can be identified or not. This conclusion should be based on an evaluation of the available data against the agreed ExAG.

8.5. Conclusions and recommendations

- Gimsing et al. (2019) provide a list of items to report targeted monitoring studies. This list is, however, not comprehensive. The Panel suggests creating a list comparable to the list of items to characterise publicly available monitoring data.
- The reported items should be explicitly evaluated with respect to the relevant elements of the selected ExAG to demonstrate whether the ExAG has been reached.
- An analysis of the most important uncertainties from the monitoring study must be part of the monitoring report. A balanced discussion of false positives and false negatives must be part of this analysis. The Panel suggests that guidance needs to be developed on how to use results from this uncertainty analysis in regulatory assessments.

9. General conclusions and recommendations

The Panel concludes that Gimsing et al. (2019) provide many recommendations on how to design and conduct groundwater monitoring studies without giving specific guidance on how to design, conduct and evaluate groundwater monitoring studies for regulatory purposes. However, the information provided in Gimsing et al. (2019) can serve as a basis for developing detailed and harmonised guidance for targeted monitoring studies. In this section, the most important conclusions and recommendations are listed. More detailed conclusions and recommendations are listed in the respective sections.

9.1. The exposure assessment goal

- Before detailed and harmonised guidance can be developed, risk managers must define which ecosystem services need to be protected where and when (or SPG). This definition is crucial for designing appropriate risk assessment schemes in which ecotoxicological effects can be combined with exposure measured or simulated in the field.
- The current leaching assessment scheme for groundwater aims primarily at the ecosystem service 'provisioning of clean drinking water' and does not consider impacts on organisms. However, when defining SPGs for groundwater, further relevant ecosystem services must be considered, especially those addressing groundwater as habitat for non-target species and the impact of groundwater contamination on surface water and its communities.
- The exposure assessment is an integral part of any risk assessment scheme. To define which exposure should be used to evaluate the SPG for groundwater in a structured and unambiguous way, the so-called ExAG must be defined. This allows answering questions such as (i) where, in what environmental compartment, and for what time frame should exposure be estimated, or (ii) how conservative should the exposure estimate be, i.e. what percentage of the exposure situations in the field should be covered in the risk assessment.
- The Panel recommends operationalising the ExAG in a transparent and unambiguous manner using the six elements of ExAGs as a guideline, considering also different SPGs. This operationalisation should preferably be done by risk assessors after consultation of risk managers. Important elements to decide upon by the risk managers are (i) the depth relative to the soil surface and relative to the groundwater layer, (ii) the time window over which concentrations can be averaged and (iii) the desired overall level of protection by selecting a spatiotemporal percentile of the statistical population of relevant concentrations.
- As water sampled in the uppermost groundwater is assumed to be on the conservative side with respect to concentrations over the entire depth (due to processes such as degradation during downward transport, or dilution with uncontaminated water from non-treated areas), the Panel concludes that concentrations at depths ranging from 1 to 10 m below the soil surface and that originate from the upper layer of the groundwater body (1–2 m below the groundwater table), represent well, i.e. in a conservative way, the concentration in the entire groundwater body or aquifer.

9.2. The tiered approach

- Groundwater monitoring is the highest tier (i.e. Tier-4) of the FOCUS leaching assessment scheme. According to the principles of a tiered approach, higher tiers should provide a less conservative and more realistic estimate of the target quantity. The Panel notes that well performed monitoring studies provide for a more realistic exposure assessment than the lower (modelling) tiers. If the monitoring depth is greater than 1 m (which is the evaluation depth for the models used at the lower tiers), measured concentrations will generally also be lower than the concentrations obtained in the lower tiers. This can, however, not *always* be guaranteed because processes that are not included in the models can result in higher ground concentrations. Because monitoring studies provide more realistic exposure assessments, the Panel concludes that (a well-conducted) Tier-4 study can overrule results from lower tier studies.
- Preferential flow through macropores is one of the known processes that could make the assessment at Tier-4 more conservative. The Panel recommends investigating if it is possible to include a harmonised description of preferential flow in the leaching models that are currently used in the leaching assessment and to investigate the consequences for the groundwater concentration in the aquifer. Ultimately, preferential flow could be incorporated into the lower tiers of FOCUS groundwater.

9.3. The vulnerability assessment

- The Panel agrees with Gimsing et al. (2019) that a central question in the design and interpretation of groundwater monitoring studies is that of groundwater vulnerability. Applicants must demonstrate that the selected monitoring sites represent realistic worst-case conditions as specified in the ExAG. For this purpose, it is necessary to model groundwater vulnerability in the entire area of use of the pesticide. It was demonstrated in several studies that substance properties play an important role in the vulnerability of groundwater to pesticide leaching. The consequence is that leaching models are needed to map groundwater vulnerability. Both process-based simulation models, and statistical metamodels can be used for this purpose; the Panel considers the so-called index approach inappropriate.
- The Panel recommends developing harmonised models for assessing groundwater vulnerability at EU level. This should include the development of harmonised geodata. The models and data sets should be well-documented, brought under version control and preferably be accessible through a user-friendly interface. The geodata should be regularly updated to accommodate for the effects of climate change and land-use change.
- The Panel recommends using the annual mass flux at 1-m depth as an indicator of groundwater vulnerability. Given the lack of data on deeper soil layers, and because groundwater vulnerability is to a large extent determined by processes occurring in the topsoil, the Panel considers this the only practical way forward. The Panel notes, however, that vulnerability of the groundwater aquifer may differ from the groundwater vulnerability at 1-m depth. This gives uncertainty in the vulnerability assessment, which must be considered when selecting the monitoring sites and interpreting results at a later stage.
- The Panel notes that the groundwater vulnerability assessment serves two purposes, i.e. (i) identification of potential monitoring sites, and (ii) setting monitoring studies into a larger spatial context. The examples given in Gimsing et al. (2019) are illustrative; however, the Panel notes that binding guidance for regulatory use needs to be developed. The Panel agrees with Gimsing et al. (2019) that there is currently no available data set that can be used to identify monitoring sites at *field-level* across the EU. So, the Panel proposes to clearly distinguish between these two phases of site selection, i.e. selection of vulnerable areas and/or regions, and selection of the actual monitoring sites within these vulnerable areas. In the second phase, applicants must demonstrate that the actual site characteristics (e.g. soil and climate data) are comparable to those derived from the vulnerability assessment. Criteria need to be set when these characteristics are sufficiently equal.
- Although groundwater vulnerability is substance dependent, the Panel accepts to base the selection of monitoring sites on the substance with the most important concentration exceedances. Whether the selected sites are sufficiently conservative for other substances must be demonstrated in the context setting procedure thereafter.
- The in-context setting procedure can be used to demonstrate a safe use in different FOCUS zones than the FOCUS zone(s) where the monitoring has been performed. The in-context procedure can also be used for the authorisation process at the national level, if requirements set by national authorities are met.

9.4. Study design of targeted monitoring programmes

- Gimsing et al. (2019) describe two types of targeted monitoring programmes, i.e. in-field and edge-of-field study designs. In-field study designs generally profit from the ease of proving hydrological connectivity between the monitoring wells and the treated fields, which is essential when monitoring is used for regulatory purposes. The disadvantage of in-field study designs is the risk of groundwater contamination due to the installation of the wells, rather than leaching. In this respect, edge-of-field study designs are the preferred option. To safeguard collection of maximum residues in the groundwater originating from the adjacent field, the Panel recommends collecting samples in the uppermost shallow groundwater (1- to 2-m filter screen installed near the groundwater table). This would also be in line with the proposed ExAG.
- A prerequisite for the regulatory use of monitoring data is the availability of complete data on the use history of the products containing the respective active substances. The application data must be reported in sufficient detail for the whole monitoring period and the relevant

preceding time span (considering the travel times of the substance). Scaling is needed if the actual application rate differs from the definition of uses according to the GAP. For this, guidance needs to be developed. The Panel also recommends developing guidance how to relate the actual application frequencies reported in monitoring studies to the definition of the uses according to the GAP.

- As mentioned above, it is crucial to demonstrate hydrological connectivity between the field(s) that received the application (and for which data on use is documented) and the sampling well. Connectivity can be proved using measurements of tracers (e.g. bromide) or pseudo-tracers (e.g. metabolites). An alternative is to simulate pesticide transport from the treated field to the well with a combination of a leaching model and a groundwater model (this is referred to as 'time-of-flight modelling'). The Panel considers the combined use of measurements of pseudo-tracers (e.g. metabolites) or tracers (e.g. bromide) with site-specific time-of-flight modelling exercises to be a suitable and practical approach to demonstrate connectivity.
- The Panel considers time-of-flight modelling a compulsory part of this analysis, because pseudo-tracers and tracers may have shorter travel times than the substance of concern. The Panel notes that time-of-flight modelling requires a high degree of experience with hydrological modelling and site-specific model parameterisation. Further guidance and training courses would be needed if these approaches become more important in regulatory use.
- The Panel recommends developing a statistical approach to determine the minimum number of sites required for regulatory groundwater monitoring.
- The Panel notes that site-selection involves a high workload for both regulators and applicants. A network of standardised monitoring sites could therefore help regulatory acceptance of groundwater monitoring studies. The Panel recommends investigating if such a monitoring network could be established at EU level.

9.5. Publicly available monitoring programmes

- Monitoring data collected by third party organisations, such as environmental agencies, water companies and research institutions, are designed for purposes other than for the authorisation of PPPs under Regulation (EC) 1107/2009. They provide important information on the current state of and possible trends in groundwater bodies and provide additional information on the applicability and interpretation of targeted monitoring studies due to vast amount of information that is generally available. Therefore, the Panel recommends including publicly available monitoring data in the risk assessment as supportive information. If the monitoring data shows, for example, large-scale exceedances of the drinking water limit for a substance that has passed the lower tiers of the groundwater leaching assessment, applicants could be asked to provide additional information to demonstrate the safe use of the product.
- At present, publicly available monitoring data often do not meet the main quality criteria for a Tier-4 risk assessment. The Panel recommends using these data as supportive information in a weight-of-evidence approach and not to use these data to overrule earlier tiers. To include publicly available monitoring data as supportive information in the risk assessment, the data need to comply with quality criteria. These quality criteria must be established in a future guidance document. Guidance is further needed on the interpretation of monitoring data of shared metabolites.
- Access to the data for regulators could be improved by making them available in a standardised European database, together with the metadata and procedures to establish a plausible relationship between the presence of the substance in groundwater and the authorised use.

9.6. Reporting results from groundwater monitoring programmes

- Gimsing et al. (2019) give some considerations for the study report of monitoring studies to be submitted to regulatory authorities. A list of items to be reported is included as well. This list is, however, not comprehensive for targeted monitoring studies. The Panel suggests creating a fixed list of items to be reported.
- The discussion section must contain an analysis of the most important uncertainties from the monitoring study. A balanced discussion of false positives and false negatives must be part of this analysis. The Panel suggests that guidance needs to be developed on how to use results from this uncertainty analysis in regulatory assessments.

References

Adriaanse PI, Buddendorf WB, Holterman HJ and ter Horst MMS, 2022. Supporting the development of exposure assessment scenarios for Non-Target Terrestrial Organisms to plant protection. Development of Exposure Assessment Goals. EFSA supporting publication 2022:EN-7661. https://doi.org/10.2903/sp.efsa.2022.EN-7661

Allen R and Walker A, 1987. The influence of soil properties on the rates of degradation of metamitron, metazachlor and metribuzin. Pesticide Science, 18(2), 95–111. https://doi.org/10.1002/ps.2780180204

Annable MD, Hatfield K, Cho J, Klammler H, Parker BL, Cherry JA and Rao PSC, 2005. Field-scale evaluation of the passive flux meter for simultaneous measurement of groundwater and contaminant fluxes. Environmental Science & Technology, 39(18), 7194–7201. https://doi.org/10.1021/es050074g

Badawi N, Karan S, Haarder EB, Rosenbom AE, Gudmundsson L, Hansen CH, Nielsen CB, Plauborg F, Kørup K and Olsen P, 2022. The Danish pesticide leaching assessment programme. Monitoring Results May 1999–June 2020. Geological Survey of Denmark and Greenland. Available online: www.plap.dk

Boesten JJTI and van der Linden AMA, 1991. Modelling the Influence of Sorption and Transformation on Pesticide Leaching and Persistence. Journal of Environmental Quality, 20(2), 425–435. https://doi.org/10.2134/jeq1991.00472425002000020015x

Boesten JJTI, van der Linden AMA, Beltman WHJ and Pol JW, 2015. Leaching of plant protection products and their transformation products. Proposals for improving the assessment of leaching to groundwater in the Netherlands. Alterra, Wageningen, Report 2264. Available online: https://edepot.wur.nl/222949

Boesten JJTI, 2018. Conceptual considerations on exposure assessment goals for aquatic pesticide risks at EU level. Pest Management Science, 74(2), 264–274. https://doi.org/10.1002/ps.4701

Brandi-Dohrn FM, Dick RP, Hess M and Selker JS, 1996. Suction cup sampler bias in leaching characterization of an undisturbed field soil. Water Resources Research, 32(5), 1173–1182. https://doi.org/10.1029/96WR00290

Bronswijk JJB and Evers-Vermeer JJ, 1990. Shrinkage of Dutch clay soil aggregates. Netherlands Journal of Agricultural Science, 38(2), 175–194. https://doi.org/10.18174/njas.v38i2.16603

Burns M, Reichenberger S, Pires J and Tripault H, 2015. How representative are the Northern groundwater scenarios of the actual conditions in the Northern zone? Report of a study commissioned by the Danish Environmental Protection Agency and financed by the Nordic Council of Ministers. v. 1.1, 124 pp.

Castano-Sanchez A, Hose GC and Reboleira A, 2020. Ecotoxicological effects of anthropogenic stressors in subterranean organisms: a review. Chemosphere, 244, 125422. https://doi.org/10.1016/j.chemosphere.2019.125422

Danielopol DL, Pospisil P and Rouch R, 2000. Biodiversity in groundwater: a large-scale view. Trends in Ecology and Evolution, 15(6), 223–224.

De Jonge H, de Jonge LW and Jacobse OH, 2000. [^{14}C]Glyphosate transport in undisturbed topsoil columns. Pest Management Science, 56(10), 909–915. https://doi.org/10.1002/1526-4998(200010)56:10%3C909::AID-PS227%3E3.0.CO;2-5

De Sousa L, van den Berg F and Heuvelink GBM, 2022. A soil organic matter map for arable land in the EU. Wageningen Environmental Research report 3126. https://doi.org/10.18174/556312

De Weerd H and Leijnse A, 1997. Assessment of the effect of kinetics on colloid facilitated radionuclide transport in porous media. Journal of Contaminant Hydrology, 26(1–4), 245–256. https://doi.org/10.1016/S0169-7722(96)00072-1

Di Lorenzo T, Di Marzio WD, Sáenz ME, Baratti M, Dedonno AA, Iannucci A, Cannicci S, Messana G and Galassi DMP, 2014. Sensitivity of hypogean and epigean freshwater copepods to agricultural pollutants. Environmental Science and Pollution Research Internal, 21(6), 4643–4655. https://doi.org/10.1007/s11356-013-2390-6

Di Lorenzo T, Borgoni R, Ambrosini R, Cifoni M, Galassi DMP and Petitta M, 2015. Occurrence of volatile organic compounds in shallow alluvial aquifers of a Mediterranean region: baseline scenario and ecological implications. Science of the Total Environment, 538, 712–723. https://doi.org/10.1016/j.scitotenv.2015.08.077

Dolan T, Howsam P, Parsons DJ and Whelan MJ, 2013. Is the EU drinking water directive standard for pesticides in drinking water consistent with the precautionary principle? Environmental Science Technology, 47(10), 4999–5006. https://doi.org/10.1021/es304955g

EC (European Commission), 2009. Guidance on groundwater status and trend assessment. Guidance document no 18, Publications Office, 2012. https://doi.org/10.2779/77587

EC (European Commission), 2014. Assessing potential for movements of active substances and their metabolites to ground waters in the EU. Report of the FOCUS Ground Water Work Group. EC document reference Sanco/13144/2010 v3, 613 pp. Available online: https://esdac.jrc.ec.europa.eu/public_path/projects_data/focus/gw/NewDocs/focusGWReportOct2014.pdf

EC (European Commission), 2017. Guidance on monitoring and surveying of impacts of pesticide use on human health and the environment under Article 7(3) of Directive 2009/128/EC establishing a framework for Community action to achieve the sustainable use of pesticides (referred to as the Sustainable Use Directive). Commission notice of 10.10.2017. EC (2017), 6766 final. Available online: https://food.ec.europa.eu/system/files/2017-10/pesticides_sup_monitoring-guidance_en.pdf

EC (European Commission), 2022. Proposal for a directive of the European parliament and of the council amending directive 2000/60/EC establishing a framework for community action in the field of water policy, directive 2006/118/EC on the protection of groundwater against pollution and deterioration and Directive 2008/105/EC on environmental quality standards in the field of water policy. COM (2022) 540 final. Available online: https://environment.ec.europa.eu/publications/proposal-amending-water-directives_en

EFSA (European Food Safety Authority), 2017. EFSA Guidance Document for predicting environmental concentrations of active substances of plant protection products and transformation products of these active substances in soil. EFSA Journal 2017;15(10):4982. https://doi.org/10.2903/j.efsa.2017.4982

EFSA (European Food Safety Authority), 2022. Revised guidance on the risk assessment of plant protection products on bees (Apis mellifera, Bombus spp. and solitary bees) Public consultation number PC-0217 (In preparation).

EFSA (European Food Safety Authority), 2023. Peer review of the pesticide risk assessment of the active substance S-metolachlor excluding the assessment of the endocrine disrupting properties. EFSA Journal 2023;21(2):7852. https://doi.org/10.2903/j.efsa.2023.7852

EFSA FEEDAP Panel (EFSA Panel on Additives and Products or Substances used in Animal Feed), Bampidis V, Bastos M, Christensen H, Dusemund B, Kouba M, Kos Durjava M, Lopez-Alonso M, Lopez Puente S, Marcon F, Mayo B, Pechova A, Petkova M, Ramos F, Sanz Y, Villa RE, Woutersen R, Brock T, de Knecht J, Kolar B, van Beelen P, Padovani L, Tarres-Call J, Vettori MV, Azimonti G, 2019. Guidance on the assessment of the safety of feed additives for the environment. EFSA Journal 2019;17(4):5648, 78 pp. https://doi.org/10.2903/j.efsa.2019.5648

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. https://doi.org/10.2903/j.efsa.2010.1821

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2012. Scientific opinion on the science behind the guidance for scenario selection and scenario parameterisation for predicting environmental concentrations of plant protection products in soil. EFSA Journal 2012;10(2):2562, 76 pp. https://doi.org/10.2903/j.efsa.2012.2562

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. https://doi.org/10.2903/j.efsa.2013.3290

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2013b. Scientific opinion on the report of the FOCUS groundwater working group (FOCUS, 2009): assessment of lower tiers. EFSA Journal 2013;11(2):3114, 29 pp. https://doi.org/10.2903/j.efsa.2013.3114

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2014. Scientific Opinion on good modelling practice in the context of mechanistic effect models for risk assessment of plant protection products. EFSA Journal 2014;12(3):3589. https://doi.org/10.2903/j.efsa.2014.3589

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Ockleford C, Adriaanse PI, Berny P, Brock T, Duquesne S, Grilli S, Hernandez-Jerez AF, Bennekou SH, Klein M, Kuhl T, Laskowski R, Machera K, Pelkonen O, Pieper S, Stemmer M, Sundh I, Teodorovic I, Tiktak A, Topping CJ, Wolterink G, Craig P, de Jong F, Manachini B, Sousa P, Swarowsky K, Auteri D, Arena M and Rob S, 2017. Scientific Opinion addressing the state of the science on risk assessment of plant protection products for in-soil organisms. EFSA Journal 2017;15(2):4690. https://doi.org/10.2903/j.efsa.2017.4690

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2015. Scientific Opinion addressing the state of the science on risk assessment of plant protection products for non-target arthropods. EFSA Journal 2015;13(2):3996. https://doi.org/10.2903/j.efsa.2015.3996

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Ockleford C, Adriaanse P, Berny P, Brock T, Duquesne S, Grilli S, Hernandez-Jerez AF, Bennekou SH, Klein M, Kuhl T, Laskowski R, Machera K, Pelkonen O, Pieper S, Stemmer M, Sundh I, Teodorovic I, Tiktak A, Topping CJ, Wolterink G, Aldrich A, Berg C, Ortiz-Santaliestra M, Weir S, Streissl F, Smith RH, 2018. Scientific Opinion on the state of the science on pesticide risk assessment for amphibians and reptiles. EFSA Journal 2018;16(2): 5125. https://doi.org/10.2903/j.efsa.2018.5125

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Hernandez-Jerez AF, Adriaanse PI, Aldrich A, Berny P, Duquesne S, Focks A, Marinovich M, Millet M, Pelkonen O, Pieper S, Tiktak A, Topping CJ, Widenfalk A, Wilks M, Wolterink G, Binaglia M, Chiusolo A, Serafimova R, Terron A and Coja T, 2022. Scientific opinion on toxicity of pyrethroid common metabolites. EFSA Journal 2022;20(10):7582. https://doi.org/10.2903/j.efsa.2022.7582

EFSA SC (Scientific Committee), 2016. Guidance to develop specific protection goals options for environmental risk assessment at EFSA, in relation to biodiversity and ecosystem services. EFSA Journal 2016;14(6), e04499. https://doi.org/10.2903/j.efsa.2016.4499

EFSA SC (Scientific Committee), 2018. Guidance on Uncertainty Analysis in Scientific Assessments. EFSA Journal 2018;16(1):5123, https://doi.org/10.2903/j.efsa.2018.5123

EMA (European Medicines Agency), 2018. Guideline on assessing the environmental and human health risks of veterinary medicinal products in groundwater. EMA/CVMP/ERA/103555/2015, 13 pp. Available online: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-assessing-environmental-human-health-risks-veterinary-medicinal-products-groundwater_en.pdf

Farsad A, Herbert SJ, Hashemi M and Sadeghpour A, 2012. An automated suction lysimeter for improved soil water sampling. Vadose Zone Journal, 11(4), vzj2012.0003. https://doi.org/10.2136/vzj2012.0003

Flury M, 1996. Experimental evidence of transport of pesticides through field soils: a review. Journal of Environmental Quality, 25(1), 25–45. https://doi.org/10.2134/jeq1996.00472425002500010005x

FOCUS (FOrum for the Co-ordination of pesticide fate models and their USe), 2000. FOCUS groundwater scenarios in the EU review of active substances v.2.2 – report of the FOCUS Groundwater Scenarios Workgroup. EC Document Reference Sanco/321/2000 rev.2, 202 pp. Available online: https://esdac.jrc.ec.europa.eu/public_path/projects_data/focus/gw/docs/FOCUS_GW_Report_Main.pdf

Gimsing AL, Agert J, Baran N, Boivin A, Ferrari F, Gibson R, Hammond L, Hegler F, Jones RL, König W, Kreuger J, van der Linden AMA, Liss D, Loiseau L, Massey A, Miles B, Monrozies L, Newcombe A, Poot A, Reeves GL, Reichenberger S, Rosenbom AE, Staudenmaier H, Sur R, Schwen A, Stemmer M, Tüting W and Ulrich U, 2019. Conducting groundwater monitoring studies in Europe for pesticide active substances and their metabolites in the context of Regulation (EC) 1107/2009. Journal of Consumer Protection and Food Safety, 14(Suppl 1), S1–S93. https://doi.org/10.1007/s00003-019-01211-x

Helsel D, 2012. Statistics for Censored Environmental Data Using Minitab and R. 2nd Edition. John Wiley & Sons, Inc. 324 p. https://doi.org/10.1002/9781118162729

Helsel DR, 2006. Fabricating data: how substituting values for nondetects can ruin results, and what can be done about it. Chemosphere, 65(11), 2434–2439. https://doi.org/10.1016/j.chemosphere.2006.04.051

Herrmann M and Sur R, 2021. Natural attenuation along subsurface flow paths based on modelling and monitoring of a pesticide metabolite from three case studies. Environmental Sciences Europe, 33(59), 1–36. https://doi.org/10.1186/s12302-021-00490-2

Heuvelink GBM, Burgers SLGE, Tiktak A and van den Berg F, 2010. Uncertainty and stochastic sensitivity analysis of the GeoPEARL pesticide leaching model. Geoderma, 155(304), 186–192. https://doi.org/10.1016/j.geoderma.2009.07.004

Holdt G, Janzen W, König W, Osterwald A and Wöltjen J, 2021. Considering pH dependent degradation and adsorption in soil for groundwater leaching assessment. Draft version 0.1.

Kruijne R, Montforts M, Janssen G, Poot A, de Jonge M, van den Berg F and Meering M, 2023. Proposal for the use of regulator groundwater monitoring results for the authorisation of plant protection products in the Netherlands. Wageningen Environmental Research report 3217. https://doi.org/10.18174/585273

Jamin P, Dollé F, Chisala B, Orban P, Popescu I-C, Hérivaux C, Dassargues A and Brouyère S, 2012. A regional flux-based risk assessment approach for multiple contaminated sites on groundwater bodies. Journal of Contaminant Hydrology, 127(1–4), 65–75. https://doi.org/10.1016/j.jconhyd.2011.07.001

Jarvis NJ, Moeys J, Hollis JM, Reichenberger S, Lindahl AML and Dubus IG, 2009. A conceptual model of soil susceptibility to macropore flow. Vadose Zone Journal, 8(4), 902–910. https://doi.org/10.2136/vzj2008.0137

Kreft A and Zuber A, 1978. On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions. Chemical Engineering Science, 33(11), 1471–1480. https://doi.org/10.1016/0009-2509(78)85196-3

Kupfersberger H, Klammler G, Schumann A, Brückner L and Kah M, 2017. Modeling subsurface fate of S-metolachlor and metolachlor ethane sulfonic acid in the Westliches Leibnitzer Feld aquifer. Vadose Zone Journal, 17(1), 1–12. https://doi.org/10.2136/vzj2017.01.0030

Larson MJH and Jarvis NJ, 2000. Quantifying interactions between compound properties and macropore flow effects on pesticide leaching. Pest Management Science, 56(2), 133–141. https://doi.org/10.1002/(SICI)1526-4998(200002)56:2<133::AID-PS103>3.0.CO;2-N

Lindahl AML, Dubus IG and Jarvis N, 2009. Site classification to predict the abundance of the deep-burrowing earthworm *Lumbricus terrestris* L. Vadose Zone Journal, 8(4), 911–915. https://doi.org/10.2136/vzj2008.0140

Magid J and Christensen N, 1993. Soil solution sampled with and without tension in arable and heathland soils. Soil Science Society of America Journal, 57(6), 1463–1469. https://doi.org/10.2136/sssaj1993.03615995005700060011x

Małoszewski P and Zuber A, 1982. Determining the turnover time of groundwater systems with the aid of environmental tracers: 1. Models and their applicability. Journal of Hydrology, 57(3–4), 207–231. https://doi.org/10.1016/0022-1694(82)90147-0

Manenti R, Piazza B, Zhao Y, Padoa Schioppa E and Lunghi E, 2021. Conservation studies on groundwaters' pollution: challenges and perspectives for stygofauna communities. Sustainability, 2021(13), 7030. https://doi.org/10.3390/su13137030

Michel C, Baran N, André L, Charron M and Joulian C, 2021. Side effects of pesticides and metabolites in groundwater: impact on denitrification. Frontiers in Microbiology, 12, 662727. https://doi.org/10.3389/fmicb.2021.662727

Mohaupt V, Völker J, Altenburger R, Birk S, Kirst I, Kühnel D, Küster E, Semeradova S, Subelj G and Whalley C, 2020. Pesticides in European rivers, lakes and groundwaters – data assessment. ETC/ICM technical report 1/2020: European Topic Centre in Inland, Coastal and Marine Waters, 86 pp. Available online: https://www.eionet.europa.eu/etcs/etc-icm/products/etc-icm-reports/etc-icm-report-1-2020-pesticides-in-european-rivers-lakes-and-groundwaters-data-assessment

NRC (National Research Council), 2000. Natural Attenuation for Groundwater Remediation. The National Academies Press, Washington, DC. https://doi.org/10.17226/9792

Orgiazzi A, Ballabio C, Panagos P, Jones A and Fernández-Ugalde O, 2018. LUCAS Soil, the largest expandable soil dataset for Europe: a review. European Journal of Soil Science, 69, 140–153. https://doi.org/10.1111/ejss.12499

Parker JC and van Genuchten MT, 1984. Flux-averaged and volume-averaged concentrations in continuum approaches to solute transport. Water Resources Research, 20, 866–872. https://doi.org/10.1029/WR020i007p00866

Petersen CT, Jensen HE, Hansen S and Bender Koch C, 2001. Susceptibility of a sandy loam soil to preferential flow as affected by tillage. Soil and Tillage Research, 58(1–2), 81–89. https://doi.org/10.1016/S0167-1987(00)00186-0

Rosenbom AE, Kjær J, Henriksen T, Ullum M and Olsen P, 2009. Ability of the MACRO model to predict long-term leaching of metribuzin and diketometribuzin. Environmental Science Technology, 43(9), 3221–3226. https://doi.org/10.1021/es802752x

Rosenbom AE, Olsen P, Plauborg F, Grant F, Juhler RK, Brüsch W and Kjær J, 2015. Pesticide leaching through sandy and loamy fields – long-term lessons learnt from the Danish Pesticide Leaching Assessment Programme. Environmental Pollution, 201(June 2015), 75–90. https://doi.org/10.1016/j.envpol.2015.03.002

Roth K and Jury WA, 1993. Linear transport models for adsorbing solutes. Water Resources Research, 29(4), 1195–1203. https://doi.org/10.1029/92WR02537

Ruqin FAN, Xiaoping Z, Xueming Y, Aizhen L, Shuxia JIA and Xuewen C, 2013. Effects of tillage management on infiltration and preferential flow in a black soil. Northeast China. Chinese Geographical Science, 23(3), 312–320. https://doi.org/10.1007/s11769-013-0606-9

Sanders EC, Abou Najm MR, Mohtar RH, Kladivko E and Schulze D, 2012. Field method for separating the contribution of surface connected preferential flow pathways from flow through the soil matrix. Water Resources Research, 48, W04534. https://doi.org/10.1029/2011WR011103

Scorza Júnior RP and Boesten JJTI, 2005. Simulation of pesticide leaching in cracking clay soil with the PEARL model. Pest Management Science, 61(5), 432–448. https://doi.org/10.1002/ps.1004

Scorza Júnior RP, Jarvis NJ, Boesten JJTI, van der Zee SEATM and Roulier S, 2007. Testing MACRO (version 5.1) for pesticide leaching in a Dutch clay soil. Pest Management Science, 63(10), 1011–1025. https://doi.org/10.1002/ps.1434

Scorza Júnior RP, Smelt JH, Boesten JJTI, Hendriks RFA and van der Zee SEATM, 2004. Preferential Flow of Bromide, Bentazon, and Imidacloprid in a Dutch Clay Soil. Journal of Environmental Quality, 33(4), 1473–1486.

Shoari N and Dubé J-S, 2018. Toward improved analysis of concentration data: embracing nondetects. Environmental Toxicology and Chemistry, 37(3), 643–656. https://doi.org/10.1002/etc.4046

Sket B, 1999a. The nature of biodiversity in hypogean waters and how it is endangered. Biodiversity & Conservation., 8, 1319–1338. https://doi.org/10.1023/A:1008916601121

Sket B, 1999b. High biodiversity in hypogean waters and its endangerment: the situation in Slovenia, the Dinaric Karst, and Europe. Crustaceana, 72(8), 767–779. Available online: http://www.jstor.org/stable/20106199

Sposito G and Barry DA, 1987. On the Dagan model of solute transport in groundwater: foundational aspects. Water Resources Research, 23(10), 1867–1875. https://doi.org/10.1029/WR023i010p01867

Syngenta, 2014. A9396G, S-Metolachlor 960 g/L EC, Notification of an active substance under Commission Regulation (EU) 844/2012, Document M-CP, Section 9, Fate and Behaviour in the Environment, 21 November 2014 updated 1 August 2016, A9396G_11212.

Tiktak A, Boesten JJTI, Egsmose M, Gardi C, Klein M and Vanderborght J, 2013. European scenarios for exposure of soil organisms to pesticides. Journal of Environmental Science and Health, Part B, 48(9), 703–716. https://doi.org/10.1080/03601234.2013.780525

Tiktak A, Hendriks RFA and Boesten JJTI, 2012a. Simulation of movement of pesticides towards drains with a preferential flow version of PEARL. Pest Management Science, 68(2), 290–302. https://doi.org/10.1002/ps.2262

Tiktak A, Hendriks RFA, Boesten JJTI and van der Linden AMA, 2012b. A spatially distributed model of pesticide movement in Dutch macroporous soils. Journal of Hydrology, 2012(470–471), 316–327. https://doi.org/10.1016/j.jhydrol.2012.09.025

Tiktak A, Poot A, Jene B, Ghafoor A, van den Berg F, Hoogeweg G, Klein M, Stemmer M, Sur R and Sweeney P, 2020. Spatially distributed leaching modelling of pesticides in the context of Regulation (EC) 1107/2009. Problem definition document. Available online: https://esdac.jrc.ec.europa.eu/projects/sdlm

Tiktak A, van der Linden AMA and Boesten JJTI, 2003. The GeoPEARL model. Description, applications and manual. RIVM report 716601007/2003. Available online: https://www.rivm.nl/publicaties/geopearl-model-description-applications-and-manual

Tiktak A, van der Linden AMA and Uffink G, 2005. Pesticide transport in the groundwater at the national scale: coupling an unsaturated zone model with a groundwater flow model. In: Thomson NR (ed.). Bringing Groundwater Quality Research to the Catchment Scale. Volume 297. IAHS Publication. pp. 441–448.

Tiktak A, van den Berg F and Poot A, 2022. The Dutch decision tree for the evaluation of the leaching of plant protection products, Revised 2022 version. RIVM-report 2022-0048. https://doi.org/10.21945/RIVM-2022-0048

Vanclooster M, Boesten JJTI, Trevisan M, Brown CD, Capri E, Eklo OM, Gottesbüren B, Gouy V and van der Linden AMA, 2000. A European ring test of pesticide-leaching models: methodology and major recommendations. Agricultural Water Management, 44(1–3), 1–19. https://doi.org/10.1016/S0378-3774(99)00081-5

Van den Berg F, Tiktak A, Heuvelink SLGE, Brus DJ, de Vries F, Stolte J and Kroes JG, 2012. Propagation of uncertainties in soil and pesticide properties to pesticide leaching. Journal of Environmental Quality, 41(1), 253–261. https://doi.org/10.2134/jeq2011.0167

Vanderborght J, Tiktak A, Boesten JJTI and Vereecken H, 2011. Effect of pesticide fate parameters and their uncertainty on the selection of 'worst-case' scenarios of pesticide leaching to groundwater. Pest Management Science, 67(3), 294–306. https://doi.org/10.1002/ps.2066

Vanderborght J and Vereecken H, 2001. Analyses of locally measured bromide breakthrough curves from a natural gradient tracer experiment at Krauthausen. Journal of Contaminant Hydrology, 48(12), 23–43. https://doi.org/10.1016/S0169-7722(00)00176-5

Vanderborght J and Vereecken H, 2007. Review of dispersivities for transport modelling in soils. Vadose Zone Journal, 6(1), 29–52. https://doi.org/10.2136/vzj2006.0096

Van der Linden AMA, Boesten JJTI, Cornelese AA, Kruijne R, Leistra M, Linders JBHJ, Pol JW, Tiktak A and Verschoor AJ, 2004. The new decision tree for the evaluation of pesticide leaching from soils. RIVM report number 601450019.

Van der Linden AMA, Tiktak A, Boesten JJTI and Leijnse A, 2009. Influence of pH-dependent sorption and transformation on simulated pesticide leaching. Science of the Total Environment, 407(10), 3415–3420. https://doi.org/10.1016/j.scitotenv.2009.01.059

Van der Salm C, van den Toorn A, Chardon WJ and Koopmans GF, 2012. Water and nutrient transport on a heavy clay soil in a fluvial plain in the Netherlands. Journal of Environmental Quality, 41(2012), 229–241. https://doi.org/10.2134/jeq2011.0292

Van Schaik NLMB, Hendriks RFA and van Dam JC, 2010. Parameterization of macropore flow using dye-tracer infiltration patterns in the SWAP model. Vadose Zone Journal, 9(1), 95–106. https://doi.org/10.2136/vzj2009.0031

Van Stiphout TPJ, van Lanen HAJ, Boersma OH and Bouma J, 1987. The effect of bypass flow and internal catchment of rain on the water regime in a clay loam grassland soil. Journal of Hydrology, 95(1–2), 1–11. https://doi.org/10.1016/0022-1694(87)90111-9

Vereecken H, Schnepf A, Hopmans JW, Javaux M, Or D, Roose T, Vanderborght J, Young MH, Amelung W, Aitkenhead M, Allison SD, Assouline S, Baveye P, Berli M, Brüggemann N, Finke P, Flury M, Gaiser T, Govers G, Ghezzehei T, Hallett P, Hendricks Franssen HJ, Heppell J, Horn R, Huisman JA, Jacques D, Jonard F, Kollet S, Lafolie F, Lamorski K, Leitner D, McBratney A, Minasny B, Montzka C, Nowak W, Pachepsky Y, Padarian J, Romano N, Roth K, Rothfuss Y, Rowe EC, Schwen S, Šimůnek J, Tiktak A, van Dam J, van der Zee SEATM, Vogel HJ, Vrugt JA, Wöhling T and Young IM, 2016. Modeling soil processes: review, key challenges, and new perspectives. Vadose Zone Journal, 15(5), 1–57. https://doi.org/10.2136/vzj2015.09.0131

Villholth KG, Jensen KH and Fredericia J, 1998. Flow and transport processes in a macroporous subsurface-drained glacial till soil. I: field investigations. Journal of Hydrology, 207(1–2), 98–120. https://doi.org/10.1016/S0022-1694(98)00129-2

Villholth KG, Jarvis NJ, Jacobsen OH and de Jonge H, 2000. Field investigations and modelling of particle-facilitated pesticide transport in macroporous soil. Journal of Environmental Quality, 29(4), 12981309. https://doi.org/10.2134/jeq2000.00472425002900040037x

Walker A and Thompson JA, 1977. The degradation of simazine, linuron and propyzamide in different soils. Weed Research, 17(6), 399–405. https://doi.org/10.1111/j.1365-3180.1977.tb00500.x

Zaki MH, Moran D and Harris D, 1982. Pesticides in groundwater: the aldicarb story in Suffolk County, NY. American Journal of Public Health, 72(12), 1391–1395. https://doi.org/10.2105/ajph.72.12.1391

Zhang Y, Baeumer B and Benson DA, 2006. Relationship between flux and resident concentrations for anomalous dispersion. Geophysical Research Letters, 33, L18407. https://doi.org/10.1029/2006GL027251

Abbreviations

ADI	acceptable daily intake
AGES	Austrian Agency for Health and Food Safety
a.i.	active ingredient
DAR	Draft Assessment Report
EfAG	effect assessment goal
ERC	ecotoxicological relevant concentration
EREQ	ecotoxicological relevant exposure quantity

ExAG	exposure assessment goal
FOCUS	FOrum for the Co-ordination of pesticide fate models and their Use
GAP	Good Agricultural Practice
GeoPEARL	The spatially distributed version of PEARL (see below)
GLP	Good Laboratory Practice
K_{om}	organic matter–water partition coefficient
LOD	limit of detection
LOQ	limit of quantification
MACRO	leaching model, specifically developed for addressing macroporous water flow in soils
MTC	months at target concentration
PEARL	Pesticide Emission at Regional and Local Scales. A pesticide fate model intended for higher-tier exposure and leaching assessments
PEC	predicted environmental concentration
PELMO	Pesticide Leaching Model. A pesticide fate model intended for higher-tier exposure and leaching assessments
PLAP	Danish Pesticide Leaching Assessment Programme
PPR	Plant Protection Products and their Residues
PPP	plan protection product
PRZM	Pesticide Root Zone Model for calculating fate and behaviour of substances in the unsaturated zone of the soil
RC	relevant concentration
RMS	Rapporteur Member State
SC	Scientific Committee
SCoPAFF	Standing Committee of Plant Animal Food and Feed
SDLM	spatially distributed leaching model
SPG	specific protection goal
SU	spatial unit
SETAC	Society of Environmental Toxicology and Chemistry
UBA	Umweltbundesamt

Appendix A – Resident vs flux concentrations

A.1. Background

In a tiered approach, the same target quantity should be used throughout all its steps for matter of comparison. In the lower tiers, the target quantity is the 80th percentile of 20 measures of the annual mass flux transported over the 1-m depth horizontal plane in soil, divided by the annual water volume transported across the same plane (flux concentration). In Tier-4 groundwater monitoring studies, water is sampled over a certain filter length of the groundwater well for a short period of time in the order of hours/day. Note that the principle of collection of groundwater samples from monitoring wells is comparable to the way groundwater abstraction wells for drinking water operate. Questions related to the consistency of the tiered approach that need to be addressed are (i) the type of concentration that is sampled by an extraction well (flux vs resident concentration) and how do these types relate to each other; (ii) the temporal scale of averaging (year vs hours) of the target quantity; and (iii) the spatial scale of averaging (field vs filter length) of the target quantity.

A.2. Type of concentration

Solute concentrations are expressed in mass of solute per volume of liquid phase. Concentrations are measured by sampling devices in the laboratory or in the field and appear in solute transport equations and boundary conditions of models. However, two conceptually different definitions of solute concentration exist, namely resident and flux concentrations, depending on the mode used for averaging. Kreft and Zuber (1978) defined resident concentration as the mass of solute per unit volume of fluid contained in an elementary volume of the system at a given instant. The flux concentration was defined as the mass of solute per unit volume of fluid passing through a given cross section at an elementary time interval and equals the ratio of the solute mass flux density to the volumetric water flux density for one dimensional flow. Note that the flux concentration is not defined when the water flux density equals zero.

Sampling devices in porous media dictate the type of concentration that is being measured. Soil coring, for example, provides a measure of the total mass of solute at a given time, i.e. the instantaneous mass density of solute, in a fixed volume of the porous medium. This type of concentration fits to the definition of the resident concentration for conservative tracers, like bromide, which reside in the liquid phase. In the case of pesticides, the total resident concentration may be partitioned over several compartments in the bulk soil volume with the aid of process models (Roth and Jury, 1993), e.g. partitioning of solutes over the liquid and the solid phase assuming equilibrium or time-dependent sorption. In this case, the solute flux density is associated with dissolved solutes in the mobile phase only.

The flux concentration is a flow-weighted concentration, which would be measured by a perfect passive sampling device that does not interfere with the native flow and allows the solution to enter the device at the same rate as the flow rate in the undisturbed porous media. Examples of sampling devices that collect in theory water samples over a defined cross-sectional area with minimal disturbance of the native flow are passive capillary samplers for the vadose zone (Brandi-Dohrn et al., 1996) and the passive flux meters for the saturated zone (Annable et al., 2005). The passive capillary sampler is the wick pan lysimeter. The wetted wicks function as a hanging water column and apply a suction in the range of 0 to 50 cm depending on the water flow. For the passive flux meter, a sorptive, permeable material is placed in a borehole or in a monitoring well to intercept the contaminated groundwater for a certain period. Thereafter, the sorptive material is removed for extraction and analysis. Thus, passive samplers would be the least ambiguous devices to sample flux concentrations.

Unfortunately, measurements of flux concentration are usually not possible without disturbing the native water flow, especially when suction is applied to extract water. Moreover, suction can be applied at different modes, continuous at constant or time-variable pressure levels or intermittent as a falling head. A flux concentration would be collected as the rate of water extraction approaches the soil water or groundwater flow rate, whereas a resident concentration would be sampled as the rate of extraction becomes infinite (Sposito and Barry, 1987). Thus, suction-based solution sampling devices will provide concentration measurements that approach either a flux or a resident concentration.

Suction devices measure resident concentration when they instantaneously extract all water in the immediate vicinity of the sampling location within a known volume. Suction devices, like suction cups that operate in a falling head mode, approximate resident concentrations (Magid and Christensen, 1993; Brandi-Dohrn et al., 1996), where the time for soil water extraction is short compared to the sampling intervals. In this mode the larger pores empty first, and then water is extracted from the smaller pores.

Suction devices measure flux concentration when they extract water at the ambient flow velocity through a known cross-sectional area and a defined period without disturbing the water flow field. Suction devices that operate in a continuous mode with constant or time-variable suction approximates flux concentration. The least disturbance of the water flow field, and thus the closest resemblance to a flux concentration, will occur when the applied suction equals the time-variable ambient pressure head in the soil. An example of such device in the vadose zone is the so-called automated suction lysimeter as described by Farsad et al. (2012). In this device, a time-variable suction is applied to a ceramic pressure plate, which complies to the suction measured by a tensiometer which is placed in the close vicinity of the plate.

Solution samplers in groundwater observation wells operated by suction are most appropriately viewed as (local) flux concentrations (e.g. Parker and van Genuchten, 1984). Although the time for extraction is short compared to the time interval between measurements, pores in the aquifer will, in contrast to the unsaturated zone, not be emptied when suction is applied to extract water. Water from the larger pores or from regions with higher hydraulic conductivity will contribute more to the sample than water from smaller pore or from regions with lower hydraulic conductivity, owing to a higher resistance to water flow. Thus, water from areas with higher flow velocities are preferably, i.e. flux proportional, sampled. Furthermore, the rate of water extraction by the sampler will resemble the groundwater flow rate (flux concentration) closer than an infinite rate (resident concentration). Note that the installation of the well and the applied suction will disturb the water flow field to some instance. The magnitude of these disturbances will affect the closeness of resemblance of the sampled concentration to a flux concentration.

To conclude, although solution-sampling devices may approach exact measures of either resident or flux concentrations, it is more likely that the concentration measurement will lie somewhere in between. Nevertheless, depending on the measuring device, the measured concentration is assigned to be either a resident or a flux concentration for the interpretation. In that case, solution samplers in groundwater observation wells operated by suction most appropriately reflect local flux concentrations.

A.3. Comparison between flux and resident concentration

Although the relationship between flux and resident concentration may be extended to three spatial dimensions and transient flow conditions (Kreft and Zuber, 1978; Parker and van Genuchten, 1984), we restrict the discussion here to one-dimensional, steady state flow conditions. Then, the flux (C^f) and resident (C_l^r) concentration in the liquid phase are related by (Kreft and Zuber, 1978):

$$C^f = C_l^r - \lambda \frac{\partial C_l^r}{\partial z},$$

where λ is dispersivity and z is depth. Dispersivity, which is the ratio of the dispersion coefficient and the pore-water velocity, reflects the typical spatial scale where variations in microscopic pore-water velocities are averaged out. The flux concentration is equal to the resident concentration if the dispersivity is zero or if the gradient of the resident concentration is zero around the depth of interest. Zero dispersivity means that variations in pore-water velocities at the microscopic scale tend to zero (Parker and van Genuchten, 1984), which resembles piston flow. Piston-like flow conditions also implies that transport by molecular diffusion over a given cross-section is negligible. Małoszewski and Zuber (1982) phrased that flux and resident concentration are unequal, 'whenever flow lines of different velocities have different solute contents', i.e. deviate from piston flow, or 'when there is a gradient of concentration along the flow lines', i.e. $\partial C^r/\partial z \neq 0$. In porous media with very high Peclet numbers, i.e. the ratio of the characteristic transport length to the dispersivity is large, flux and resident concentrations will give similar estimates for the transport behaviour. Note that the characteristic transport length is defined at the scale of the surface of application or of the point of injection to the reference plane and not locally.

Zhang et al. (2006) explored functional relationships between flux and (mobile and total) resident concentrations for non-Fickian (anomalous) dispersion behaviour, i.e. for preferential flow conditions and/or for the existence of diffusion-limited immobile zones, using a random walk particle tracking model. When diffusion-limited immobile zones exist, the total resident concentration is composed of the resident concentration in the mobile and immobile phase. For porous media with preferential flow and/or diffusion-limited immobile zones, they showed that the flux concentration at the point of observation, which is located at a certain distance from the point of injection or application, is larger than the resident concentrations for very early arrival times. This has also been shown experimentally by Brandi-Dohrn et al. (1996). The total resident concentration now resembles the resident concentration in the mobile phase, since the solute particles reside in the mobile phase only, with minimal interaction of zones with lower flow velocities. In the long term, most solute particles will reside in the immobile or less mobile regions due to diffusion-like processes. Now, the total resident concentration is higher than the resident concentration in the mobile phase. The latter approximates the flux concentration, since any mobile solute particles are driven mainly by advection in the mobile phase and are the only particles that can cross the point of observation within a unit time.

Deviations between flux concentrations and the resident concentrations in the mobile phase diminish over time in porous systems with the presence of mobile and immobile zones. This implies that deviations between flux concentrations and the resident concentrations in the mobile phase will be less pronounced with increasing distance from the point of application and the point of observation, i.e. for increasing transport times. For groundwater monitoring wells, this means that deviations between flux and resident concentrations will be smaller with increasing the filter depth or increasing distance between the points of application and extraction. Also, the importance of preferential transport, inherently advection-dominated, will be less important at larger transport times or transport distances. Conversely, the difference in type of concentration can be decisive for the interpretation for short transport times or transport distances.

For illustrative purposes, Figure A.1 shows a comparison between the flux concentration and the resident concentration in the liquid phase at 1-m depth for the FOCUS Hamburg scenario calculated by PEARL v4.4.4 and PELMO v5.5.3. Pesticide D was applied annually at a rate of 1 kg/ha in winter cereals 1 day after emergence. Both types of concentrations are similar for the PEARL simulation, except for a few peak events. The hints at the long-term behaviour as pointed out in Zhang et al. (2006). Note that PEARL simulates a downward and upward water flux density, and concurrent mass flux density, at the 1-m depth according to the hydraulic gradients imposed at the upper boundary by precipitation/irrigation and evapotranspiration. This transient up- and downward movement of solute particles at the depth of interest may also reduce the importance of advective transport. On the other hand, the flux concentration is always higher than the resident concentrations in the liquid phase in the PELMO simulations, hinting at a more advection-dominated transport (short-term behaviour in Zhang et al., 2006). Indeed, PELMO is not able to simulate upward flow at the 1-m depth, with only downward directed solute and mass flux densities when the soil compartments are at field capacity (generally not during summertime).

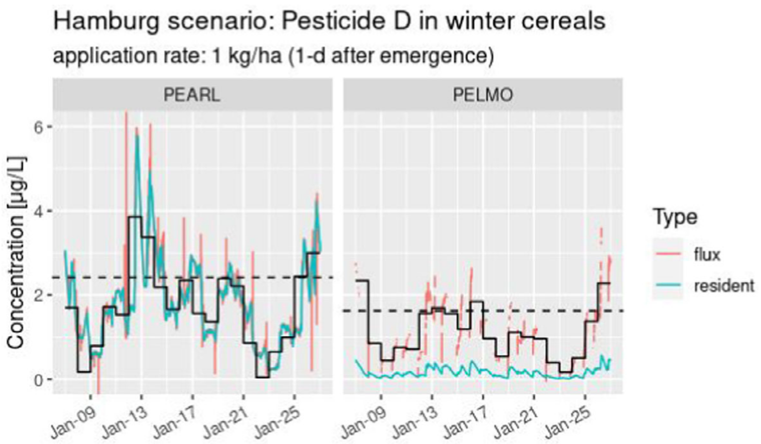

Figure A.1: Comparison between the daily flux concentration and resident concentration in the liquid phase of pesticide D at 1-m depth in the Hamburg scenario calculated by PEARL v4.4.4 and PELMO v5.5.3. Pesticide D was applied annually in winter cereals 1 day after first emergence at a rate of 1 kg/ha. The first 6 years are excluded (warming up phase). The solid black line represents the annual averaged flux concentration, and the dashed black line is the 80th percentile of the 20 annually averaged flux concentrations. Note that the flux concentration is not defined in case the water flux density equals zero, as is often the case for the PELMO model

To conclude, the discrepancy between the flux and resident concentration is dictated by the magnitude of the ratio between dispersion and mean advection (dispersivity) and mechanisms of solute retention. The discrepancy increases with the increasing dispersivity and with the existence of preferential pathways and/or diffusion-limited immobile zones, especially for short transport times or short transport distances.

A.4. Temporal scale for averaging the target quantities

The time scales for averaging differ between the lower tiers (year or the 80th percentile of a 20-year period) and the Tier-4 groundwater monitoring samples (hours). If we assume that the target quantity is identical, it is inherent that single daily-averaged concentrations are sometimes higher than the yearly-averaged concentration, thus invalidating the tiered approach. Figure 2 gives an impression of the variability of the time scales used for averaging the flux concentrations at a daily basis (red solid line), a yearly basis (black solid line), and the 80th percentile of a 20-year period (black dashed line) at the reference depth of 1 m for the FOCUS Hamburg scenario.

To be consistent with the tiered approach, the Panel recommends that the sampling strategy of groundwater monitoring well and the proposed averaging method should reflect and be representative for the annual-averaged concentration. This sampling strategy should be continued for several years.

A.5. Area-averaged vs local flux concentration

The type of the target concentration in Tier-4 and in the lower tiers is the same, i.e. a flux concentration. However, an area-averaged flux concentration is implicitly assumed in the lower tiers (the groundwater spatial unit is the groundwater below an agricultural field of 1 ha), whereas samples from monitoring wells represent local flux concentrations. Due to spatial variability, single local concentrations can be higher than the average flux concentration for the entire groundwater spatial unit. Ideally, several monitoring wells should be installed to assess a spatially averaged concentration (e.g. Vanderborght and Vereecken, 2001). This, however, is seldomly the case. If multiple wells are installed under one field, the Panel recommends averaging measurements from those wells to get a better estimate of the average concentration for a spatial groundwater unit. To avoid false negatives, this should, however, only be done for wells that are connected to the field (see further Section 6).

Appendix B – Indicators for vulnerability assessment

As mentioned in Gimsing et al. (2019), different model outputs can be used to assess leaching vulnerability. In one of the illustrative cases, annual mass fluxes were used as an indicator of groundwater vulnerability in the context of monitoring studies (Syngenta, 2014). They argue that flux weighted concentrations may be less suitable as indicator of groundwater vulnerability in the context of groundwater monitoring studies, because high annual leachate concentrations can be associated with minimal substance mass fluxes if the water volume flux is minimal. Since mixing of these very low mass fluxes in the upper few centimetres of the groundwater already led to a huge decrease in the concentration it can be assumed that leachate concentrations that are associated with a relevant volume flux are of higher importance for the groundwater quality. Similarly, the selection of sites in statistical monitoring studies may require an aggregation of potential leachate to a larger scale. Mass fluxes can be meaningfully added or averaged whereas it is more difficult to do this with concentrations.

Comment 89 is also related to this issue: 'The approach and rationale behind using annual median mass fluxes instead of the soil pore concentration for a comparison of relative leaching risk between regions appears plausible (in example I). Nevertheless, we do not fully agree with the reasoning that the concentration is not particularly useful in the context of comparing one area with another to estimate leaching potentials. We think that for the identification of the most vulnerable sites in terms of leaching the influence of the heterogeneous distribution of pore water volume due to varying soil types and weather conditions is the most realistic reproduction of environmental conditions. Therefore, the possible differences in the outcome of both methods should be compared and scientifically evaluated'.

The main question is:

- what is the better indicator of groundwater vulnerability, annual averaged flux concentration or annual mass flux?

Other issues related to the indicator of groundwater vulnerability are:

- what percentile of the annual averaged values should be taken (median, 80th percentile, or another statistic)?
- is porewater volume (transport volume), as the most realistic reproduction of environmental conditions (soil and weather), a better indicator of the identification of the most vulnerable sites in terms of leaching?

B.1. Indicators for vulnerability assessment

In the lower tiers, the regulatory groundwater threshold value is the flux-weighted annual average concentration across a unit cross-sectional area perpendicular through the main direction of flow at 1-m depth below the soil surface (the relevant concentration). This concentration is defined as the ratio of the annual mass flux and the annual water flux. It is plausible that the extent of groundwater deterioration by pesticides is positively correlated to the pesticide mass that enters the system. Thus, the annual flux-weighted concentration and mass fluxes come into question as an indicator of leaching vulnerability into the groundwater.

For illustration, we define two hypothetical agricultural fields A and B with different soil properties, reflecting a sandy and a loamy texture, respectively. The annual water fluxes (net groundwater recharge) are 0.500 and 0.050 m, for fields A and B, respectively, due to the more pronounced capillary barrier effect in field A. These agricultural fields had received pesticide applications, which resulted in a higher annual mass flux of the pesticide in field A (5 mg/m^2) compared to field B (0.5 mg/m^2), due to a lower microbial activity in field A. Although more pesticide mass entered the groundwater from field A (5 mg) compared to field B (0.5 mg), the flux-weighted annual average concentrations are identical (10 mg/m^3). Thus, the flux-weighted annual average flux concentration is not a good indicator of the mass entering the groundwater.

Therefore, the evaluation of the impact of pesticide applications from agricultural fields on the quality of groundwater resource requires the quantification of the mass flux to the groundwater and through the groundwater to the groundwater monitoring wells (e.g. Jamin et al., 2012). The percentile of the annual mass flux in time should be defined for the vulnerability indicator in a guidance document. An 80th percentile in time would be in line with the lower tiers.

With the proposal of annual mass flux as target quantity for the indicator of leaching vulnerability, there is a difference with the target quantity used in the lower tiers, which is the annual flux weighted concentrations. At first glance, this suggests an inconsistency in the tiered approach. However, it should be noted that the annual mass flux in a vulnerability analysis is an indicator of leaching vulnerability and not a regulatory endpoint. Therefore, there is no need for a direct comparison between the target quantities used for the lower tiers and the indicator of leaching vulnerability. Furthermore, numerical simulations based on a limited number of combinations of soils and weather defined by the FOCUS scenarios showed a strong correlation between the annual mass flux and the annual flux averaged concentration (Section B.2).

B.2. Example calculations with the FOCUS groundwater models

We present an illustrative example using the nine FOCUS scenarios. PEARL v4.4.4 and PELMO v.5.5.3 were run to simulate an annual application of pesticide D in winter cereals. Pesticide D was applied 1 day after first emergence at a rate of 1 kg/ha. The 20-year timeseries of the water percolated below the target depth [mm], the substance leached the below target depth [kg/ha] and the average substance concentration in water at the target depth (μg/L) were taken from the summary files for further process. The target depth was at 1-m below the soil surface. Figure B.1 presents an overview of the simulation results.

Due to transient nature of the upper boundary condition, i.e. daily variations in precipitation and evapotranspiration over time, a downward and upward water flow will establish through the soil profile, also at the target depth of 1 m below the soil surface. This physical behaviour of downward and upward water flow at the target depth is described by PEARL, but not by PELMO. On an annual basis, this results in a net downward (positive values) or upward (negative value) movement of both water and pesticide mass across the target depth. Figure B.2 shows that all four combinations exist in the nine FOCUS scenarios of winter wheat and pesticide D. For PEARL, most data points reside in the first quadrant (net downward movement of water and substance) and least in the second quadrant (net upward movement of water combined with a net downward of substance). For PELMO, there is no upward water flow at 1-m depth below the soil surface. A net annual upward water flow and/or mass flux does not pose a leaching risk. Nevertheless, guidance is needed how to handle these cases. Note that in the FOCUS scenarios, zero is assigned to the annual flux-averaged concentrations, when there is either a net upward water movement, or a net upward substance movement, or both, i.e. to the data pairs that reside in the second, third and fourth quadrant of Figure B.2.

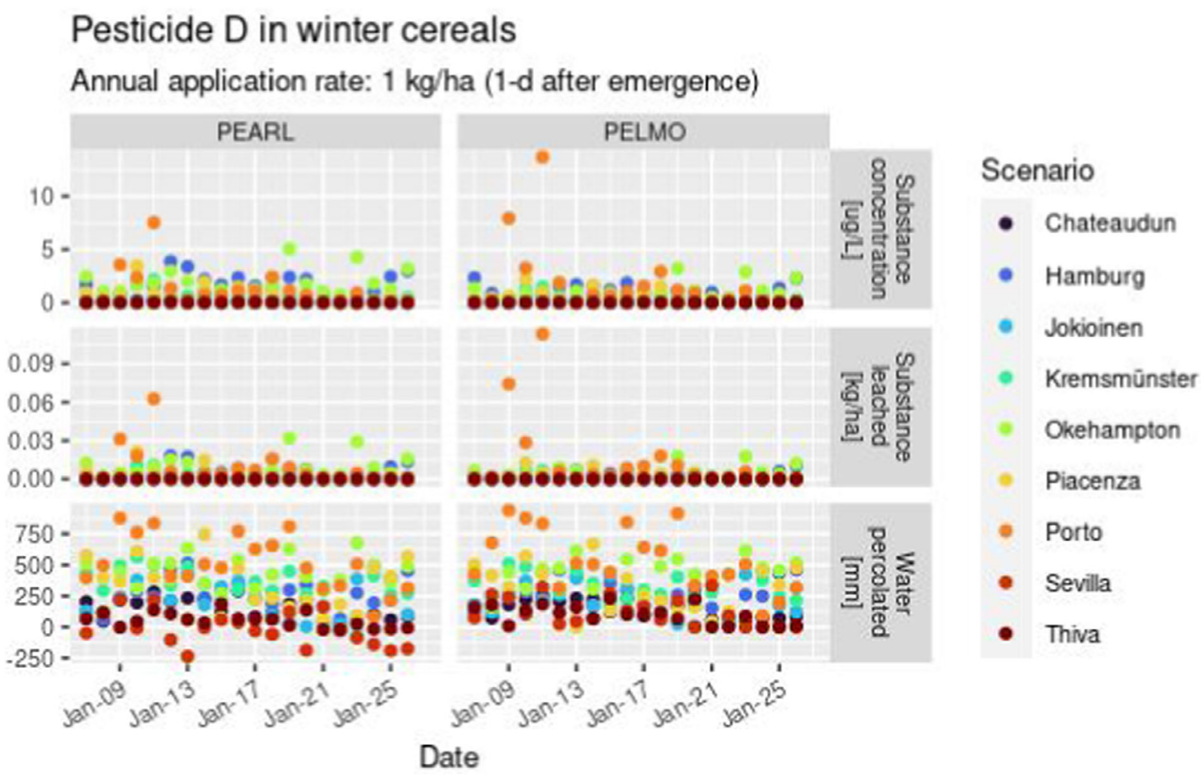

Figure B.1: Simulated annual averages of water percolated and substance leached below the 1-m depth and the substance concentration at the 1-m depth for the nine FOCUS scenarios (EC, 2014)

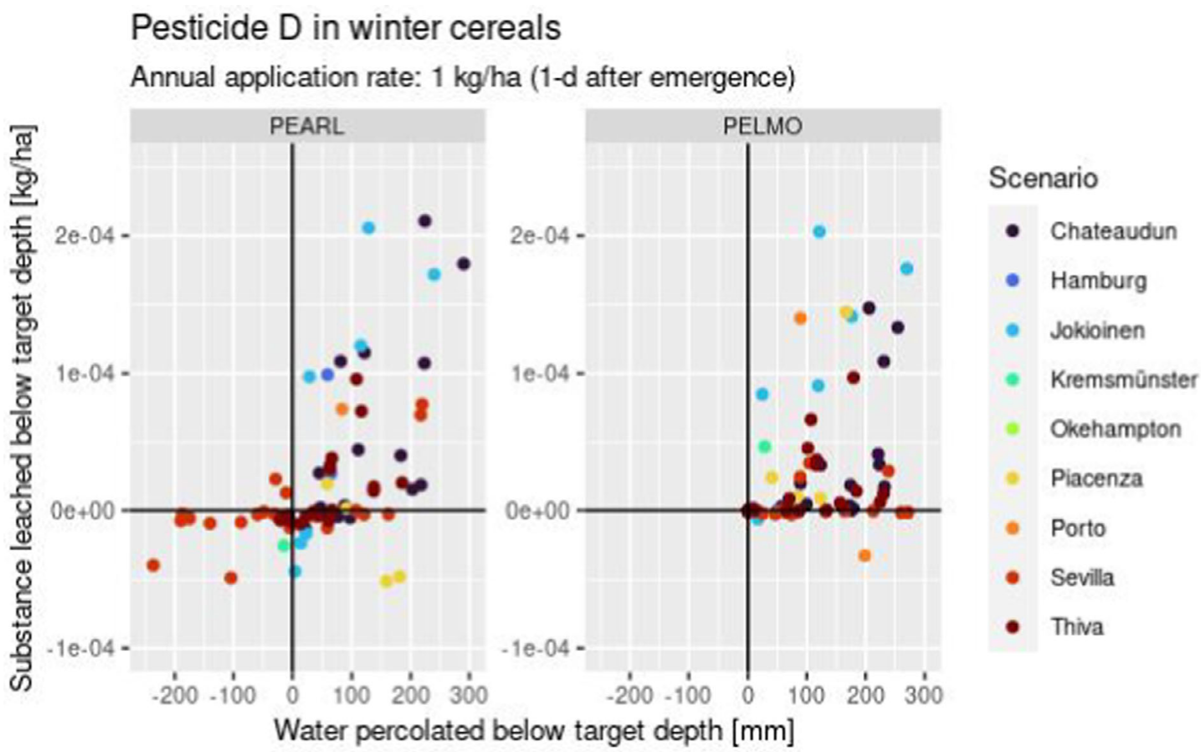

Figure B.2: Relationship between the water percolated and substance leached below target depth of 1 m. Note that both axes are truncated to have a detailed look at the values around the origin. The scenarios mentioned are the FOCUS Groundwater scenarios in EC (2014)

Figures B.3–B.5 show the correlation between the annual flux-weighted concentration, the annual water flux, and the annual mass flux for the nine FOCUS scenarios. In this example, the two possible indicators for vulnerability, i.e. the annual flux-weighted concentration and the annual mass flux, are highly correlated (correlation coefficient > 0.90 for both models). Vulnerability maps of annual mass fluxes and flux-weighted annual average concentrations (at a certain percentile) will slightly differ, especially for cases with a high annual mass flux in combination with a very low annual water flux, which yields a high flux-weighted annual average concentration. However, high flux-weighted annual average concentrations due to very low annual water flux is not represented in this example based on a limited number of combinations of soils and weather (see Figure B.5).

Very low annual water fluxes may lead to high flux-weighted annual average concentrations. A line of argumentation is that these very low water fluxes will mix in the upper few centimetres of the groundwater, which will lead to a huge dilution of the concentration in the incoming water volume. Although very low annual water fluxes did not occur in our limited simulations, Figure B.6 gives an impression of the annual dilution factor in different the scenarios. In these calculations, we assumed that the volume of groundwater recharge (net downward water flux) is completely mixed with the volume of water present in a static groundwater reservoir of 1-m depth with a porosity of 0.25 cm^3 cm^{-3}.

A stakeholder questions whether the annual mass flux is a better indicator of the identification of the most vulnerable sites in terms of leaching than the pore water volume, which we interpret as the transport volume, which they regard as the most realistic reproduction of environmental conditions (soil and weather). In the vulnerability analysis, many combinations of soil types and weather are considered in the simulation. Therefore, the heterogeneity of the transport volume in space and time is considered in these simulations and thus reflected in the annual mass fluxes.

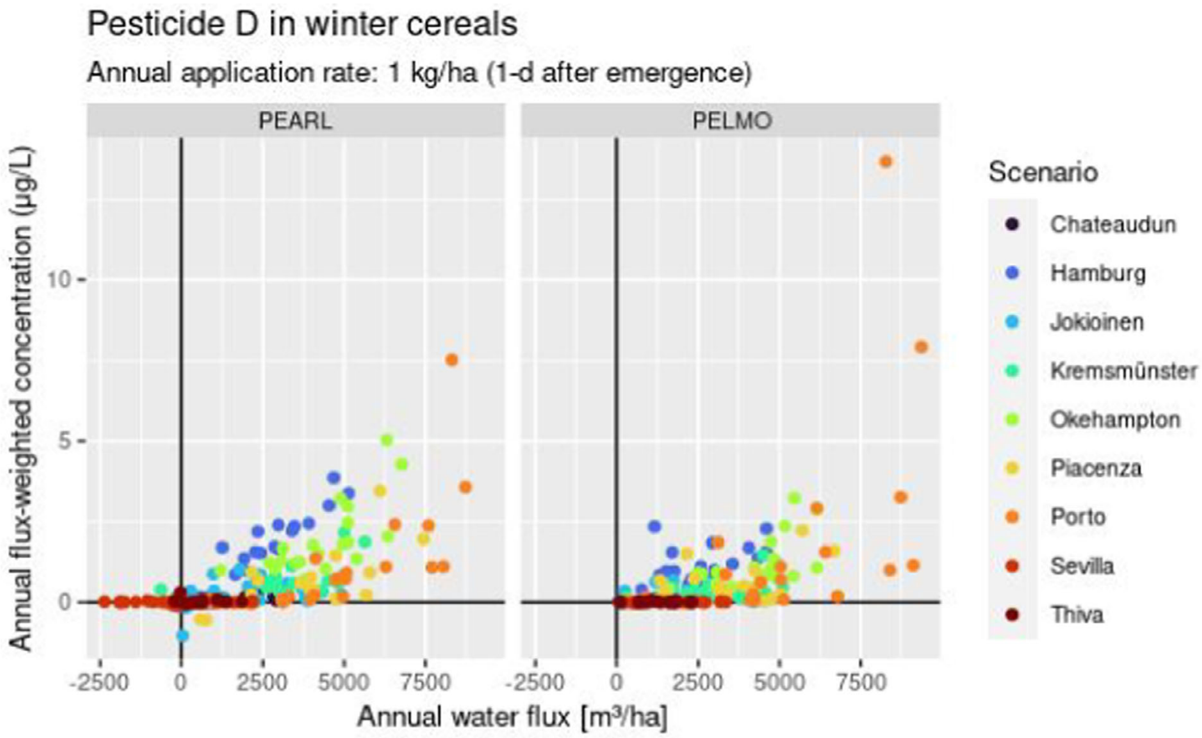

Figure B.3: Annual flux-weighted concentration as function of the annual water flux for all individual years in the nine FOCUS scenarios (EC, 2014)

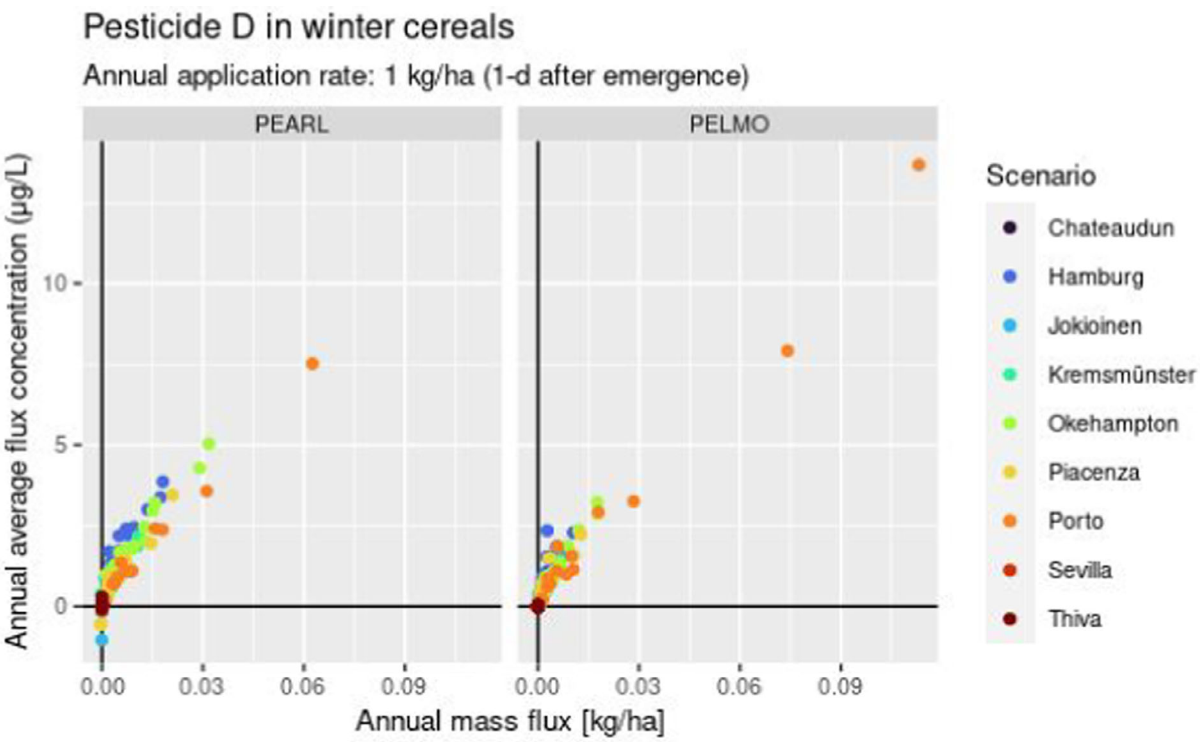

Figure B.4: Annual flux-weighted concentration as function of the annual mass flux for all individual years in the nine FOCUS scenarios (EC, 2014)

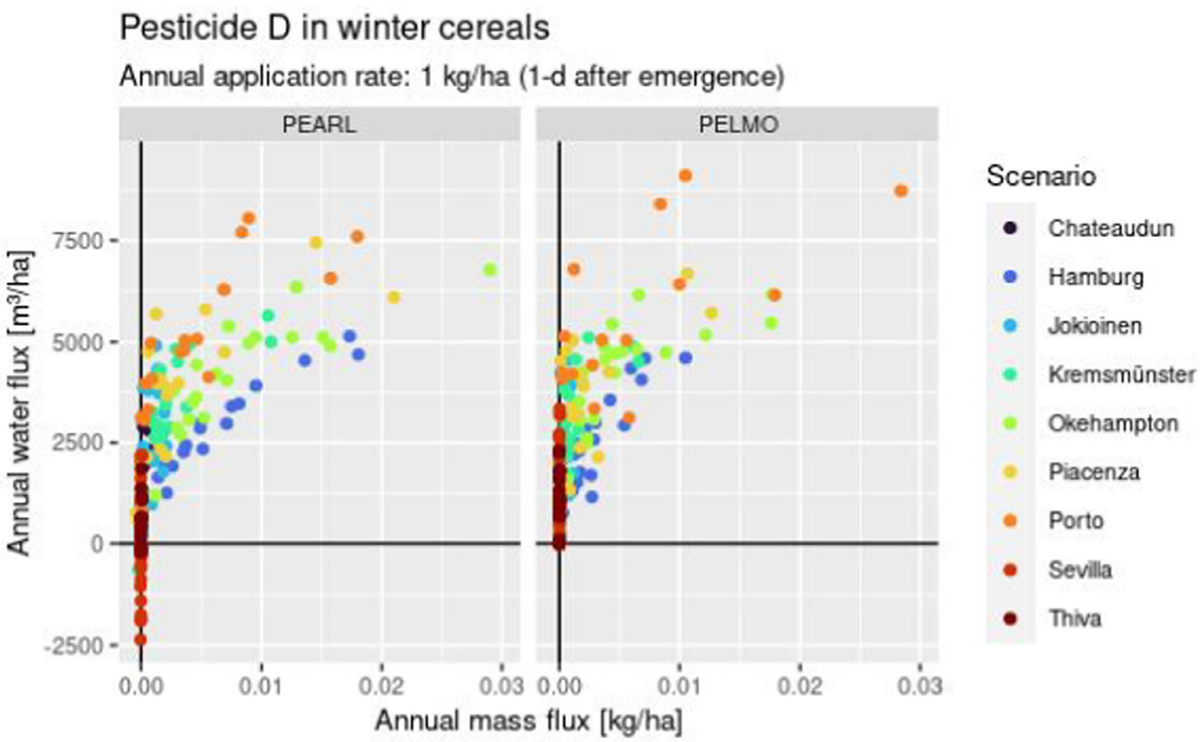

Figure B.5: Annual water flux as function of the annual water flux for all individual years in the nine FOCUS scenarios (EC, 2014)

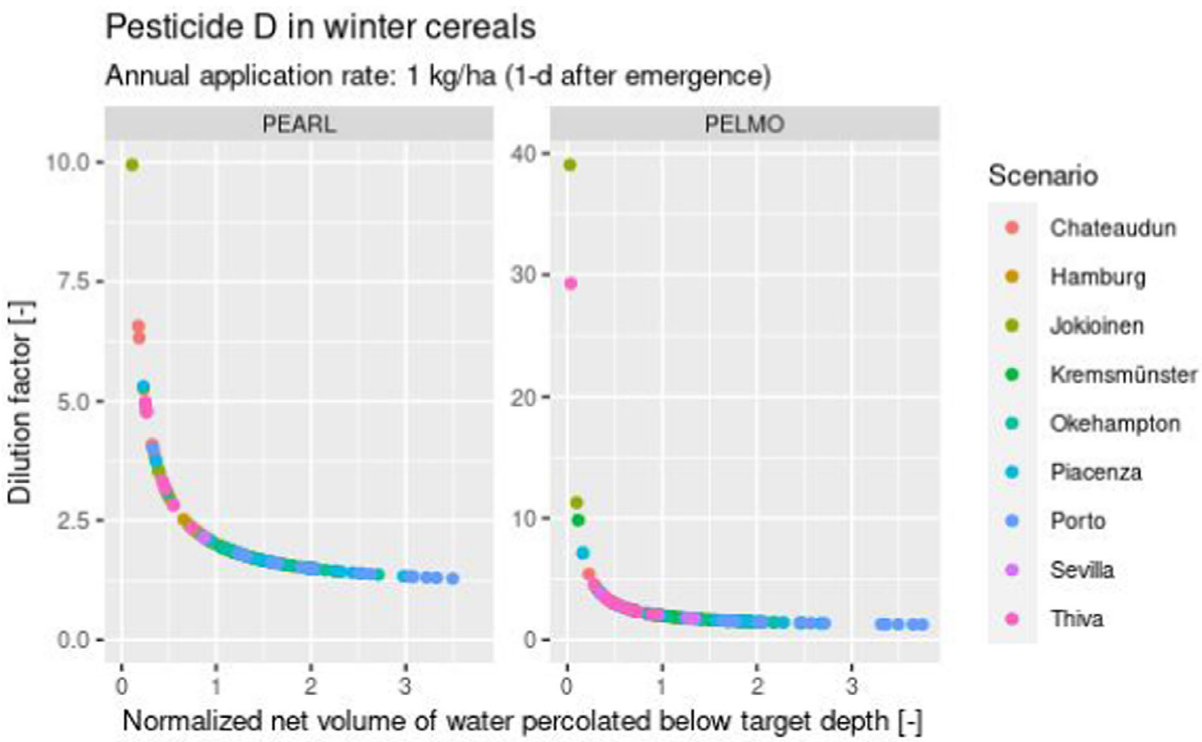

Figure B.6: Dilution factor of the concentration in the annual net volume of water percolated below the target depth. The static groundwater reservoir contains a water volume reflecting a 1-m thick aquifer layer with a porosity of 0.25 $cm^3\,cm^{-3}$

B.3. Recommendation

The Panel recommends using the annual mass flux as indicator of groundwater vulnerability.

Appendix C – List of items to characterise publicly available monitoring data[21]

As described in Section 7 of this Statement, results from publicly available monitoring programmes can provide important information on the current state and possible trends of concentrations of active substances and their metabolites from plant protection products (PPP) in groundwater in the context of Regulation (EC) 1107/2009. Monitoring data may be used to support the applicability of results from groundwater modelling or targeted monitoring studies over a longer period and/or a wider range of environmental conditions. It is therefore recommended to use the public monitoring data alongside standard modelling and/or targeted monitoring studies and interpret their results in relation to the use history of plant protection products containing a specific active substance and the applicable exposure assessment goal. However, accompanying information describing the publicly available monitoring data is often not available or is provided with different levels of detail. Poor quality of the accompanying information can complicate and bias data evaluations and limit the interpretation and comparison of publicly available monitoring results. Besides, technical data reporting for Europe provides evidence of a rather low level of harmonisation between national groundwater monitoring programmes and the need for a common data collection and storage in Europe (Mohaupt et al., 2020).

Consequently, it is necessary to sufficiently characterise the publicly available monitoring data to advance its usability in a regulatory context. Source, objective, methodology, spatial and temporal dimension of the monitoring programme should be evaluated as good as possible. The European Groundwater Monitoring Network for Regulators provided an elaborate list of items based on the list from Gimsing et al. (2019). The Panel recommends using this list for characterisation of publicly available monitoring data of active substances and/or metabolites.

The list of items in Table C.1 is aligned to groundwater monitoring data when they are provided for active substance approvals in the EU and/or zonal and national authorisations. In those cases, groundwater monitoring data are often provided for one single active substance and its metabolites only. Additionally, the list of items can support national monitoring reporting strategies of groundwater quality. In such cases, groundwater monitoring data are often provided by environmental agencies for several active substances and their metabolites.

[21] This list was established by the European Groundwater Monitoring Network for Regulators on February 7, 2022 and is endorsed by the PPR Panel. The authors of the list are Wolfram König, Emilie Farama, Roy Kasteel, Anton Poot and Signe Bonde Rasmussen.

efsa JOURNAL

Table C.1: List of items recommended to characterise publicly available groundwater monitoring data of active substances and/or metabolites for EU active substance approvals and zonal & national authorisations in context of Regulation (EC) 1107/2009 and/or for reports produced by environmental agencies or similar public institutions

Source and objective of data

List of items	Description
Authority, institution, company	Provide the name and contact of responsible authority, institution, company.
Are data publicly available or protected?	Provide a link, if original data are available.
Objective of monitoring	Provide information regarding the context of the monitoring programmes from which the data are provided, e.g. surveillance monitoring of water quality, specific monitoring programme, research context.
Date of extraction from database	Provide information about when the data used have been extracted from the database.

Characterisation of well network

List of items	Description
Monitoring area	Description of the area covered by the well network (e.g. name of member state, federal states, province, size of area).
Number of wells in the network	Provide the number of wells in the network which have been sampled. Are wells equally distributed or are certain regions in the whole area better represented?
Type of wells	Provide information about the number or percentage of different well types, e.g. observation wells, drinking water pumping stations, springs.
Aquifer types	Provide information about the different aquifer types which are covered by the sampled wells or the well network, e.g. porous, unconsolidated sediments, consolidated sediments, fractured rocks, karst aquifers.
Groundwater storey sampled	Provide information if the 1st, 2nd, 3rd, . . . groundwater storey (or aquifer layer) was sampled.
Land uses represented in the monitoring area/well network	Provide information on the land uses, e.g. arable land, pasture, permanent crops, forest, rural and urban settlements, and their relative distribution in the monitoring area/represented by the well network.
Filter screen depths below the soil surface from all wells	Provide information about the filter screen depths below the soil surface of the sampled wells or the well network, e.g. the range in depths of the centre of the filter screens.
Filter screen length from all wells	Provide information about the filter screen lengths (top level and base levels) from all sampled wells or the well network. For example, the maximum range of filter screen lengths and/or the average filter screen length could be provided.
Average water table depth from all sites	Provide information about the range of average water table depth in the sampled wells or the well network.
Information on groundwater age	If available, provide information on groundwater age.

Sampling strategy/procedure	
List of items	**Description**
Time period of sampling	Provide information about the sampling period used for reporting.
Number of samplings	Provide information about the frequency of sampling: Was each well in the network sampled with the same frequency during the sampling period? If not, specify the range of samplings. If yes, provide the number of samples.
Interval of sampling	Provide information about the sampling interval during the sampling period. Specify the range of intervals in case it is not fixed. Provide the interval in case of fixed intervals.
Subset of samplings used?	Provide information in case only a subset of samples has been used for reporting. This could be the case if number and interval of samples strongly varies during the sampling period.

Analytics (compound specific)	
List of items	**Description**
Number of laboratories involved	Is the analysis conducted by one or more laboratories?
Name and CAS-no. of active substance and analysed metabolites	If detections of an active substance and/or its metabolites are reported, please provide information on which metabolites of the active substance have been analysed in the monitoring programme. It can be expected that different subsets of metabolites from the same active substance are measured in monitoring programmes in different Member States.
Number of analysed compounds	In case all pesticides with detections are reported, please provide information on how many active substances and metabolites have been measured in the original monitoring campaign. This information can be important when data/ reports are compared with monitoring programmes in other Member States. A list of all compounds analysed could be normally provided.
Analytical method(s)	Provide general information about analytical method(s) used.
Are analytical methods stereoisomer selective?	Information should be provided if available. Identifying the right compound is important for interpretation of the results.
Limit of detection (LOD)	Information should be provided, if available. The information is maybe less important for groundwater, if < 0.1 μg/L can be ensured.
Limit of quantification (LOQ)	Information should be provided, if available. The information is maybe less important for groundwater, if < 0.1 μg/L can be ensured.

Information on PPP uses	
List of items	**Description**
Is information on PPP uses and/or the history of product authorisation within the represented monitoring area (country) available?	If available, provide information of the historical use of PPPs in the monitoring area. If detailed use information is not available, provide some general information about PPP authorisations, e.g. since when authorisations have been granted or stopped or changed over time (e.g. application rates and timings), typical crop covers in the monitoring area and/or typical selling rates in the past. This kind of information could be available from other responsible authorities or databases.

Information on PPP uses	
List of items	**Description**
Restrictions of PPP uses within the represented monitoring area	Provide information if there are/have been restrictions on PPP uses for certain well types or in larger areas of the monitoring network, which could influence the monitoring results, e.g. related to drinking water surveillance, etc. Provide information if there are/have been any restrictions/risk mitigation measures on the label of PPP uses related to the active substance approval and/or national authorisations, which could influence the monitoring results, e.g. maximum frequency or maximum rate of application defined, limitation to greenhouse uses only, etc.

Reporting[(a)]	
List of items	**Description**
Total number of samples	Provide the total number of measured samples. All available analysed samples should be provided and included in the analysis. In case individual analysed samples are not available from the original data (for example because data from agency reports are used), please specify the statistic(s) behind the provided results.
Total number of wells sampled	Provide the total number of wells with analysed samples.
Number and % of samples > LOQ	All detections of an active substance and its metabolites must be reported
Number and % of samples > 0.1 μg/L	All detections of an active substance and its metabolites must be reported
Number and % of samples > 1.0 μg/L	All detections of non-relevant metabolites must be reported
Number and % of samples > 5 μg/L	All detections of non-relevant metabolites must be reported
Number and % of samples > 10 μg/L	All detections of non-relevant metabolites must be reported
Number and % of wells with samples > LOQ	All wells with detections of an active substance and its metabolites must be reported
Number and % of wells with samples > 0.1 μg/L	All wells with detections of an active substance and its metabolites must be reported
Number and % of wells with samples > 1.0 μg/L	All wells with detections of non-relevant metabolites must be reported
Number and % of wells with samples > 2.5 μg/L	All wells with detections of non-relevant metabolites must be reported
Number and % of wells with samples > 5 μg/L	All wells with detections of non-relevant metabolites must be reported

(a): Detections of an active substance and its toxicologically relevant metabolites according to SANCO/221/2000-rev.11 (2021) should be reported in two relevant classes. Detections of toxicologically non-relevant metabolites according to SANCO/221/2000-rev.11 (2021) should be reported in four relevant classes.

Appendix D – Outcome of the public consultation

EFSA was asked by the European Commission to prepare a public consultation on the scientific paper by Gimsing et al. (2019) and take the comments into consideration for preparing the PPR statement on groundwater monitoring. The scientific and technical aspects of study designs and procedures both from targeted and/or public monitoring programmes, as well as criteria on the assessment of the reliability/quality of the groundwater monitoring information were addressed in this statement, in view of using such information in the EU regulatory exposure assessment of pesticides in line with the requirements set in Commission Implementing Regulation (EU) No 283/2013. Guidance for applicants and risk assessors on best practice in this regard were provided. EFSA performed a public consultation of the scientific paper by Gimsing et al. (2019) on the design and conduct of groundwater monitoring studies supporting pesticide groundwater exposure assessment from 26 January to 8 March 2022. The EFSA Pesticide Steering Network, risk assessors, risk managers, stakeholder and the scientific community were additionally informed via EFSA email alerts about the open public consultation. This appendix presents statistics on the comments received and answers to each of the comments. These comments were considered when updating and finalising the statement.

Screening and Evaluation of the comments received

All comments received were scrutinised and subsequently tabulated with reference to the author(s) and the section of the scientific paper by Gimsing et al. (2019) to which each comment referred. The final number of comment boxes were 94. Comments submitted formally on behalf of an organisation appear with the name of the organisation. A statistical summary of the comments received is provided in Table D.1. In Table D.2, the comments to the scientific paper are provided together with the responses by the PPR Panel.

Table D.1: Comments received from stakeholders on the scientific paper by Gimsing et al. (2019)

Organisation	Country	Total
Agence Nationale de Sécurité Sanitaire de l'Alimentation de l'Environnement et du Travail (ANSES) (France)	FRA	15
Austrian Agency for Health and Food Safety GmbH, Österreichische Agentur für Gesundheit und Ernährungssicherheit GmbH (AGES)	AUT	12
German Federal Environment Agency, Umweltbundesamt (UBA)	DEU	29
CropLife Europe	BEL	37
Spanish National Institute for Agricultural and Food Research and Technology, Instituto Nacional de Investigación y Tecnologia Agraria (INIA)	ESP	1
Total Number of comments		94

Table D.2: Comments and PPR Panel responses on the scientific the paper by Gimsing et al. (2019) per chapter

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
1	AGES (Austria)	1. Introduction	General comment: As highlighted in the executive summary of Gimsing et al., 2019, 'the report provides the SETAC EMAG-Pest GW group's recommendations on study designs and study procedures'. As such, the report describes different options on how to conduct and to assess a monitoring study without giving dedicated guidance on how to evaluate such studies or assessments from a regulatory point of view. AGES believes that there is a need to develop a more targeted regulatory guidance framework (including user-friendly software tools) for conducting, assessing and evaluating monitoring studies at different assessment levels (e.g. Regulatory Zones, FOCUS Zones or Member States), paying more attention to regulatory needs. Some issues regarding monitoring studies have not been addressed in Gimsing et al. (2019): # In monitoring studies actual application rates and frequencies are usually deviating from the intended GAP use rate and frequency (in most cases actual application rates are lower and less frequent, representing typical local use conditions). In view of AGES, there is a need for clear criteria to decide if applications at given monitoring sites adequately cover the intended use rate and frequency or not, and how to proceed if this is not the case. # In view of AGES, there is a need for a dedicated overall reliability assessment for each monitoring site and, consequently, for each set of monitoring sites finally considered relevant for a certain Member State. Among others, such a reliability assessment may account for hydraulic connectivity, leaching vulnerability, application rates and patterns, depth to groundwater, etc. Finally, clear criteria are needed when to consider an individual monitoring site or a given set of monitoring sites sufficiently reliable for regulatory decision making. # There are several statistical methods on how to calculate a certain percentile. A harmonised approach is recommended. *PPR Panel: Your comments have been considered and have been commented in the statement and in later comments*.
2	CropLife Europe	1. Introduction	CropLife Europe would like to thank EFSA for giving us the opportunity to provide comments. Our general comments are included in the enclosed document due to the character limit of this box. *PPR Panel: Noted*
3	UBA (Germany)	1. Introduction	Harmonised guidance is required before GW monitoring studies can be used and accepted as the highest tier in European GW risk assessments, zonal assessments and national authorisations according to Regulation (EC) 1107/2009. This is important for regulators in the MS since these studies are already increasingly conducted and provided by applicants. The paper provides a variety of options for exposure assessment goals, study designs, vulnerability assessment, site characterisation, sampling, reporting etc. It remains vague in many aspects and focuses on different possibilities rather than proposing reliable and binding procedures. The missing harmonised protection goal for GW in Europe is probably one reason for that. For a scientific publication we consider this a good and reasonable overview. Regarding a guidance document this is not sufficient. More and concrete guidance on prerequisites, data requirements, reproducible procedures and quality criteria will be necessary to set a harmonised standard for targeted monitoring studies that are suitable to be used for highest tier risk assessment and contain all required information (please refer to our other comments in the respective chapters).

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			We think that another subsequent process is necessary to develop an EFSA guidance document which should then be implemented in the regulation of PPPs. Further guidance development should not only consider the use of higher tier monitoring studies for assessments for active substance approvals in the EU, but also provide solutions how to handle such studies for zonal and national assessments. It currently remains open how to transfer study results to other administrative regions. One example to illustrate: How can a monitoring study conducted in Italy, evaluated by Ireland and accepted to demonstrate a 'one safe use' in the EU assessment be further considered for a product registration in Germany, when Austria is responsible for the assessment in the Central Zone? *PPR Panel: The Panel agrees that more guidance is needed; this is addressed at several places of the Statement. The issue of transferring monitoring results to different regulatory zones and Member States is addressed in Chapter 4.*
4	ANSES (France)	1. Introduction	Comment on Executive summary (Page S2): The authors underlined that 'Sampling schedules should consider the expected time required for an active substance to move through the soil into groundwater, as well as expected persistence in both soil and groundwater'. However, no data requirement relative to persistence of parent and/or metabolites in groundwater is set in Regulation 1107/2009, and the existing data from water/sediment studies cannot be considered as representative for groundwater. The availability of such data could be not straightforward. PPR Panel: Article 7.2.3 of EC 283/2013 states that studies for degradation in the saturated zone shall be discussed with national authorities. The Panel notes that if degradation in the saturated zone is taken in account in the time-of-flight, the degradation experiment must be taken from material in the saturated zone and must be representative for all possible conditions in the area of use of a pesticide, e.g. from oxic to methanogenic. This is address in Chapter 4. Comment on introduction: Agrees with the authors on the importance of the protection goal for designing monitoring studies for active substances and their metabolites and for the subsequent evaluation of the data for regulatory purposes. However, no consensus on this issue exists for the time being, leading to some difficulties to state on the choice of the different exposure assessment options proposed by the authors. The exposure assessment options described in the document are proposed according to the illustrative protection goals presented in appendix I, however it should be kept in mind that the final protection goals might be different, and this might impact the exposure assessment options. *PPR Panel: The Panel agrees that a harmonised exposure assessment goal is needed before regulatory guidance can be developed. We also agree that the goals described in* Gimsing et al. (2019) *are examples. The Panel proposes as an alternative to address six elements of the ExAG. This is addressed in Chapter 2 of the Statement.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
5	AGES (Austria)	2. Use of monitoring data as a function of various exposure assessment options	AGES is aware that there is no harmonised groundwater protection goal available, neither at the level of the EU nor at the level of Member States. As already highlighted in Gimsing et al. (2019), a well-defined protection goal is inevitable for the assessment of a monitoring study. Therefore, AGES recommends that groundwater protection goals should be defined first before providing dedicated regulatory guidance on how to conduct, assess and evaluate a monitoring study. Otherwise, a wider range of different assessment options addressing different types of protection goals should be considered in a harmonised guidance framework. *PPR Panel: The Panel agrees that development of exposure assessment goals must be carried out before regulatory guidance can be developed. This issue – together with considerations – is addressed in Chapter 2 of the Statement.*
6	CropLife Europe	2. Use of monitoring data as a function of various exposure assessment options	p. S6. Different groundwater monitoring well designs reach different levels of spatial and temporal integration. The spatial scale of a monitoring programme determines among other factors the age of the sampled water, i.e. a sample from very shallow groundwater will likely integrate percolated water from a smaller time frame compared to a sample from deeper groundwater with a wide well screen. In this sense, the temporal scale of groundwater monitoring is linked to the spatial scale and should be considered when deciding about e.g. sampling frequency. Sampling schedules can take into consideration historic residue data and if not available, estimates of travel time and groundwater flow regime on already selected sites can be used to indicate appropriate sampling regimes. *PPR Panel: We agree that sampling strategy depends on e.g. the depth of monitoring. Some considerations are given in Chapter 6 of the Statement.* p. S6 It may be useful to introduce the concept of a monitoring survey, where monitoring is used to understand the level of a compound in groundwater in general e.g. like that conducted by national authorities and a monitoring study which seeks to address specific concentrations beneath individual fields to challenge estimated residues. *PPR Panel: The Statement makes a distinction between targeted monitoring studies (Chapter 6) and public monitoring studies (Chapter 7).* p. S7 Karst is currently not modelled in any FOCUS groundwater scenarios. Due to the bespoke nature of karstic areas, protection goals defined may not be universal for karst locations due to site specific conditions. *PPR Panel: Karstic areas are a special case and should not be considered in the general groundwater monitoring assessment. They should be addressed by specific analyses, see further Chapter 2.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
7	UBA (Germany)	2. Use of monitoring data as a function of various exposure assessment options	The exposure assessment options in table 1 (and sometimes in the following chapters) are good examples but not always very precisely defined regarding the depth (e.g. option 4). The parameter ´depth below ground surface´ cannot be used alone to describe the depth of an acceptable filter screen and sampling. The depth of the groundwater table (or its range because it can change over time) needs to be determined as second important parameter because it defines the depth of the unsaturated zone important for leaching and the beginning of the saturated zone important for sampling. A description of any precise exposure option should therefore always include 'residue concentration in the groundwater x m below the water table in areas with groundwater water table < y m'. PPR Panel: Chapter 2 provides more detailed considerations of the ExAG. The Panel highlights that depth is a very important parameters and recommends that this is specifically addressed by risk managers. The Panel does not agree to the statement that depth below groundwater is not a suitable indicator, it's a risk managers decision. The paper provides different exposure assessment options (protection goals) and corresponding study designs, which we appreciate as an overview. However, from a regulatory perspective we are concerned that in practice different protection goals in member states might lead to conflicts in the zonal product authorizations. Therein, zRMS assessment is usually binding for cMS and other MS. How can this be solved regarding differing protection goals among zRMS and other MS? Please also refer to our comment on the development of a guidance document on chapter 1. *PPR Panel: The Panel agrees that development of exposure assessment goals must be carried out and how to apply these at EU level and RZ/MS level before regulatory guidance can be developed. This issue – together with considerations – is addressed in Chapter 2 of the Statement.*
8	ANSES (France)	2. Use of monitoring data as a function of various exposure assessment options	Comment 1 (Page S7, table 1): The suitability of water source from resurgence (springs) to conduct a monitoring study could be discussed. Considering ease of sampling, springs were sometimes selected in available monitoring studies. An additional exposure assessment option could be included. The cons and pros of these water samplings could be further discussed. *PPR Panel: If springs are suitable monitoring sites depends on the protection goal. Currently, the main ecosystem service to be predicted is provisioning of clean drinking water. For this, springs are less relevant. They may, however, become important if biodiversity in ecosystems that depend on groundwater are to be protected. See further Chapter 2 (Section 2.1).* Comment 2: It is anticipated that for existing monitoring studies or for studies under development, all the selected sites will not necessarily correspond to the same exposure assessment option. Some recommendations for analysing the results in such case could be provided. *PPR Panel: Existing monitoring studies (public studies) are commented in Chapter 7 of the Statement. Generally, the design of monitoring studies must be in line with the ExAG.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			Comment 3: The summary table (table 1) is very useful but could be completed to avoid confusion and to better understand the different exposure assessment options proposed. In this table and in the rest of the document, it could be clarified whether the term 'depth' correspond to the depth of the well screen, the depth below the ground surface or the depth of the water table. We note that the depth of water table is discussed for option 5 only (in section 3.4) and not for the other options. From the description of the different options, options 2 and 4 seem very similar. The differences between these 2 options need to be clarified. Based on figures 3 and 4, the length of the well screen seems to have some importance, but this should be confirmed and could be further explained. The following is reported for option 4: 'Concentration in groundwater not influenced by infiltrating water from surface water bodies'. This point could be further clarified, and it could be indicated whether it is relevant for option 4 only or also for other options. Please see additional comments in the attached document. *PPR Panel: The Panel proposes to determine the ExAG based on six dimensions as described in Chapter 2. In the view of the Panel, depth is an important part of the ExAG.*
9	CropLife Europe	2.1. Exposure assessment option 1	p. S9 'Option 1 also includes drainage water from tile drain fields as an indicator of concentrations in the upper 10 cm of the water table, although such zones of saturation may be temporary'. Sampling drainage water as a surrogate for groundwater may include residues from beyond the treated field and does not compare directly with groundwater from wells which undergo dispersion, degradation and dilution processes. It is debatable whether tile drain water represents groundwater beneath a drain. Tile drain water can be regarded as surface water and does not account for dilution and mixing within an aquifer. Furthermore, groundwater monitoring carried out in drained fields might not represent a vulnerable situation for groundwater due to the lateral transport to ditches. *PPR Panel: We consider ExAG 1 to be very conservative and not in line with the tiered approach of FOCUS groundwater. This option has therefore not further been analysed in the Statement. Generally, the Panel is not sure if groundwater monitoring in tile-drained soils must be excluded. For example, clay soils are often tile-drained and because these soils may exhibit preferential flow, it is not yet sure if these soils are not worst-cases. See further Chapter 3.* p. S9 In dedicated shallow groundwater studies, daily groundwater levels should be recorded. *PPR Panel: We agree to this, see Chapter 6.*
10	CropLife Europe	2.2. Exposure assessment option 2	p. S9 'Concentration in the upper portion of groundwater originating from below treated fields but excluding groundwater shallower than 1 m below the soil surface' It needs to be acknowledged that shallow groundwater levels can fluctuate extremely rapidly and that it may not be possible to ensure that a specific portion of the groundwater is sampled with a particular screen length. In dedicated shallow groundwater studies, daily groundwater levels should be recorded.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: We agree to this, see Chapter 6. p. S10 'Such sampling programmes may miss areas with higher concentrations, typically located near the water table under vulnerable soils, underestimating the risk of leaching to ground water'. To assess the vulnerability of a monitoring program the conditions at the monitoring sites need to be set into context. The vulnerability of a monitoring site may be determined by pedo-climatic conditions as well as by application practices. It is challenging to find wells or locations where a well can be installed that meet several vulnerability criteria. *PPR Panel: First, the depth in the ExAG is a risk managers decision. However, we agree that the variability of the concentrations in time and space is larger at shallow depths, making the monitoring more challenging (Chapter 2 and 6).*
11	CropLife Europe	2.4. Exposure assessment option 4	p. S10f Depending on the composition of the landscape, groundwater flow gradient and the well screen length the area from which the sampled groundwater origins may be larger than in option 2 and samples might integrate water that percolated from non-agricultural areas as well. However, if the well is placed close to a field, and the screen includes the upper surface of the groundwater layer, connectivity with the neighbouring field can be demonstrated. *PPR Panel: Discussion on connectivity is in Chapter 6 of the Statement.* p. S11 'This option is, in practice, essentially the same as option 2 since the highest concentrations occur in shallow groundwater which would typically be located less than 10 m from the soil surface'. Option 2 is described to explicitly focus on the upper portion the groundwater (in the first 1–2 m below the water table) while option 4 does not explicitly focus on the upper portion of groundwater but would allow e.g. 8–9 m below the water table. *PPR Panel: Selection of the depth is a risk managers decision. See Chapter 2 of the Statement*
12	CropLife Europe	2.5. Exposure assessment option 5	p.S11 'However, with this design, note that samples collected from wells installed deeper than about 3–5 m below the water table are more difficult to interpret, because such groundwater may not be originating from the field but further up gradient'. Applying tracers at the same time as crop protection products may not be possible due to practical considerations and permitting restrictions. *PPR Panel: The Panel notes that application of tracers or pseudo tracers in combination with modelling is the preferred methodology. If this is not possible, connectivity can be demonstrated using different methods, but guidance needs to be developed (see further Chapter 6 of the Statement).* p. S12 Exposure options should reflect the area to protect in a member state. If certain areas receive drinking water at depths < 10 m, then this exposure assessment option is not appropriate to show that drinking water is sufficiently protected.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: The Panel agrees that monitoring should take place at depths that provide a conservative estimate of the concentration in drinking water abstraction wells, so monitoring in the uppermost groundwater is preferred (see Chapter 2). p. S12 'Monitoring can take the form of field studies to confirm that this degradation occurs before reaching 10 m below the soil surface or more traditional monitoring studies with samples collected at depths of 10 m or greater below the soil surface'. This sentence should include 'degradation and/or dispersion'. *PPR Panel: We agree that dispersion can lower the concentration (See chapters 2 and 3).* p. S12 These sample points should not be under the influence of any point sources, should be of sound construction and not under any other non-diffuse source of crop protection products. *PPR Panel: We agree to this. The concentration measured should reflect relevant uses and not include point-sources (Chapter 8).*
13	CropLife Europe	2.7. Exposure assessment option 7	p. S12 'Concentration in raw water of a drinking-water pumping station using groundwater not influenced by surface waterbodies (no bank filtration) but not older than 50 years (this age limitation is needed to avoid that too much dilution is included in the assessment)'. Considerations need to be made on how accurate groundwater age dating techniques are, especially for younger waters. Current techniques often provide a wide range of age dates which are not accurate enough for risk assessment purposes. Sampling time for GW age dating also needs to be considered, especially for fractured geologies where several pathways to the sampling location may be active. *PPR Panel: The Panel considers these type of monitoring studies less relevant, because connectivity with relevant uses is difficult to demonstrate in catchment scale monitoring. Therefore, this is not further discussed in the Statement.* p. S12 'When there is more than one well, the concentration is the average of all wells from a pumping station at a specific sampling time'. This is not explicitly mentioned for option 6, but since options 6 and 7 are described as the same except for the groundwater age, this should also apply to option 6 *PPR Panel: The Panel proposes to use the groundwater connected to an agricultural field as the groundwater spatial unit. If there are more wells, the Panel considers averaging acceptable.*
14	CropLife Europe	2.8. Conclusion	p. S13 'While this chapter focuses on the strengths and weaknesses of various monitoring approaches, all monitoring in areas of product use can be helpful in determining whether the drinking water is being protected'. Monitoring is a higher tier option according to the FOCUS groundwater assessment scheme. As higher tiers should generate more realistic results higher tiers can supersede lower tier (i.e. modelling) results. A suitably designed GW monitoring programme can provide realistic risk assessment of groundwater (e.g. publication on herbicide groundwater monitoring in Europe by McManus et al., 2021) and can address the practicalities of

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			instrumenting sites and how to represent a number of locations. This implies that monitoring using a suitable study design is not only useful, but capable of determining if groundwater is adequately protected. Tier-4 studies designed for a specific protection goal should be considered as reliable and supportable at Tier-4, and not as weight of evidence. Publicly available monitoring data may not be able to achieve the same support as a bespoke designed Tier-4 study. The protection goal concentration from a Tier-4 groundwater study should not be compared directly with modelled concentrations. Lower tiers do not account for hydrogeological processes and rely on parametrisation of models from Tier-1 and often default values which are not representative of intrinsic properties. If Tier-4 data from a suitable designed study is available, it should be used for the risk assessment. *PPR Panel: The Panel agrees that sufficiently performed groundwater monitoring studies could be used for the risk assessment at Tier-4 and can therefore overrule the lower tiers (Chapter 3). However, the monitoring studies must be in line with the specific protection goals set by regulators and be adequate. Connectivity should for example be demonstrated and it must also be shown that the travel time to the groundwater is not too long. It is also important to demonstrate that the selected monitoring sites are sufficiently protective by setting results in context. See further Chapter 4 (in context setting) and Chapter 6 (Study designs). The Panel proposes using public monitoring data as additional evidence. However, they can seldomly be used as evidence alone (Chapter 7).*
15	UBA (Germany)	2.8. Conclusion	C4: It is concluded that 'all monitoring in areas of product use can be helpful in determining whether the drinking water is being protected'. (...) '...extensive catchment or aquifer monitoring can be sufficient to demonstrate safety for drinking water...'. (p. S13) We do not fully agree with this overall simplified conclusion according to all exposure options 1 to 7 because it does not include the time aspect as crucial parameter for the interpretation of monitoring results. Groundwater recharge, leaching of PPPs and transport of PPPs in groundwater are often slow processes. The time lag between PPP uses in a catchment area and detectable residues in groundwater normally increases with increasing depth of the installed monitoring wells and can take altogether many years up to decades. That means that catchment or aquifer monitoring (options 5, 6, 7) can only be sufficient to demonstrate safety for groundwater and/or drinking water, if a certain PPP is already approved, authorised and used in a catchment for a long time. Otherwise, false negatives are measured. Consequently, the (maximum acceptable) time lag between PPP uses and detectable residues in groundwater is an essential parameter which should be considered in the discussion of possible exposure assessment options in future to interpret measured PPP residues in groundwater according to Regulation (EC) 1107/2009. *PPR Panel: We agree that groundwater monitoring studies can only overrule lower tiers if they are sufficiently well performed. This means e.g., that connectivity must be demonstrated, travels times be assessed, etc. See Chapter 6 for considerations. Public monitoring data could provide additional evidence, but they can seldomly be used as evidence alone (Chapter 7).*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			C5: In some paragraphs in several chapters, groundwater seems to be defined as drinking water only – independent from the exposure options that explicitly focus on drinking water abstraction. This wording is irritating as we assume that – in general terms – the paper addresses groundwater. Regarding following discussions and development of a guidance document this aspect should be considered. *PPR Panel: We have noted in our Statement that groundwater performs other important ecosystem services (Chapter 2).*
16	CropLife Europe	3. Representative study designs	There should be clearer statements regarding the required minimum number of sites and sampling frequency. This does not need to be prescriptive but at least should give some boundaries. *PPR Panel: This issue is addressed in Chapter 6.* When instrumenting new sites, groundwater flow direction needs to be understood via triangulation rather than relying on topography or potentiometric maps. Other long term established well networks in MS may not need to ascertain groundwater flow in the same way as a newly installed site, due to the legacy of data and understanding on groundwater flow already known. Reference/datum points should be recorded for several permanent features of a well e.g. casing, piezometer and ground level. *PPR Panel: Agreed, this is addressed in Chapter 8.* At commercial fields, well location may not be ideal to capture all groundwater flow and locations can be compromised to satisfy landowner needs. *PPR Panel: This comment is noted.* Field leaching studies are excluded from this document, however there are several references and comparison made. There are even sections where there are statements that some of the example studies are field leaching studies. Of course, there is some gliding scale between field leaching studies and edge of field studies. But given that it is recommended that less sites are required for a field leaching study it would be good to say where the boundary is and describe this in a separate section. E.g. several wells, GLP Test Item application, tracer, but does it need pore water concentrations? Arbitrary screen lengths may not reflect the groundwater table to be sampled and the screen section should be installed to capture as much groundwater variation as possible to ensure samples can be obtained. Borehole logs identify permeable material and the groundwater strike zone which should guide the screen depth and length. *PPR Panel: The Panel agrees that the boundary between field leaching studies and groundwater monitoring studies is not always clear. However, it is highly relevant because in Tier-4 monitoring studies the measured concentration is the regulatory endpoint. The Panel recommends clarifying this difference in future guidance (Chapter 6).*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
17	UBA (Germany)	3. Representative study designs	C6: We are concerned that the implementation of higher tier monitoring studies will significantly increase the complexity of groundwater risk assessments. We expect a substantial increase of the workload for regulators. We have doubts that the current capacities in most authorities in European member states are designed for this. Therefore, any further development of guidance following the SETAC recommendations needs to bring forward solutions to reduce complexity for the conduct and assessment of targeted monitoring studies according to the Regulation (EC) 1107/2009. A better standardisation/harmonisation is required. One reason for the expected undesirable high complexity in future assessments is the new and repeated site/well selection for each single monitoring study recommended by the SETAC paper. The authors provide good reasons for that; however, most capacities are usually spent for a suitable site/well selection and sometimes well installation and hydrological site description, before the monitoring campaign with sampling, analysis and result evaluations can start. Besides, repeated site selection for each study can lead to endless discussions between applicants and regulators. Therefore, we strongly recommend reducing complexity by investing more scientific effort to develop a suitable and fixed targeted monitoring system with fixed sites and installed wells which could be repeatedly used across Europe. Such a development needs to be strengthened outside of active substance approval assessments and/or zonal and national authorisations. The PLAP system in Denmark designed as targeted monitoring system for PPP uses is a good national example which shows that fixed monitoring sites in an edge of field monitoring design can be used over several years and for several monitoring cases. We made comparable experiences with some post registration studies conducted in Germany. Complexity was reduced when the same wells were used for more than one monitoring study. *PPR Panel: The Panel agrees that fixed monitoring networks can increase regulatory acceptance of monitoring studies. This issue is address in Chapter 6 of the Statement.*
18	ANSES (France)	3. Representative study designs	Except for exposure assessments 1, 2, 3 and 4, there is no discussion or recommendation about duration of the study and sampling frequency. Although it is agreed that these are dependent on several factors, at least some indications could be provided. *PPR Panel: We addressed this issue in Chapter 2 and Chapter 6.*
19	CropLife Europe	3.1. In-field study designs for exposure assessment options 1, 2, 3, and 4	In-field designs: (page 14) Infield study designs: can be done but this would be rather atypical these days due to concerns of creating preferred pathways. It is mentioned that in-field studies should be prospective but there is reference to retrospective from time to time. A retrospective in-field study would be probably not be appropriate these days as there would be discussion whether the peak was missed. For monitoring studies in the EU, the use of in-field monitoring wells is rather problematic. Even if great care is taken with the well installation there remains the risk of creating hydraulic shortcuts or introducing topsoil containing residues to the gravel pack or saturated zone. With the very low groundwater trigger concentrations in the EU this creates a high risk of detections that may be due to the installation rather than leaching, complicating the interpretation of the results. *PPR Panel: We addressed the issue in Chapter 6.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			Sampling frequency: Quarterly or longer sampling intervals are appropriate if travel times through the unsaturated zone are 1–2 years. It is questionable if this would be appropriate, may be at a later point once the peak concentrations had reached the GW. Later in the text it is then referred to monthly sampling. This is confusing. *PPR Panel: This issue is addressed in Section 6.3.* Composite samples: mixing of water from several wells – this should not be recommended as averaging results from individual wells is also not recommended. *PPR Panel: In general, the Panel agrees that water from several wells should not be mixed. An exception is water from wells that are hydrologically connected to the same spatial unit (see Chapter 2).* Section on in-field monitoring is very vague and seems to contradict itself in several sections. The difference to field leaching studies is unclear. Ideally this section would be rewritten. *PPR Panel: Noted.*
20	UBA (Germany)	3.1.1. General study outline	Duration and sampling/Use of tracer/Determining connectivity: It is mentioned that the length of a study and the sampling interval depend on different factors like mobility and persistence of the active substance or metabolite, hydrological site characteristics, climate conditions, depth to groundwater, filter screen depth, application time, prospective and retrospective study design and the exposure assessment goal. Consequently, no clear recommendations are given in this chapter, because it depends on these factors. Monthly sampling is recommended in situations where there is limited knowledge about the hydrogeological regime in the unsaturated and the saturated zones. From our experience with targeted monitoring studies, precise information on the groundwater flow direction over time and the hydrological conditions in the unsaturated and the saturated zone at a chosen site is usually limited at the beginning of the monitoring period. This makes it difficult to estimate connectivity for a certain active substance or metabolite in a certain time frame during the study. Applicants sometimes used modelling for estimations, which is difficult without having very precise site-specific information. Therefore, most effort should be invested for determining connectivity during study duration, because it is a fundamental prerequisite that monitoring results can be interpreted according to one or few previous applications of PPP. We therefore are on the opinion that constant measurements of the groundwater flow direction as well as tracer experiments should become more important for targeted monitoring studies in the future. Besides, it seems important that new studies at new sites normally start with a higher sampling frequency, e.g. monthly sampling or even with shorter intervals to consider uncertainties in hydrological knowledge. *PPR Panel: The Panel agrees that final guidance should include recommendations about sampling frequency, repetitions, etc. See also comment #21.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
21	ANSES (France)	3.1.1. General study outline	Page S15–S16: Even when there is a good knowledge about the hydrogeological regime of the saturated zone, in our opinion a sampling effort (monthly) should be done at least for the 1st years of a monitoring study, to better understand the temporal variations. Sampling interval can then be adapted depending on results from these 1st years, after discussion with the regulatory authority. Such approach is very helpful when interpreting the results of the monitoring for each site and for sites comparison. Annual sampling is mentioned as a possibility. It should be clarified in which cases (specific type of aquifer?) and for which exposure assessment option annual sampling could be considered as relevant and sufficient. *PPR Panel: Agreed. Often, applicants provide time-of-flight-modelling exercises to estimate the sampling frequency required. However, as the modelling assumptions may not be able to consider all processes in the real world, the sampling schedule of monitoring studies must guarantee capturing seasonal fluctuations in concentrations of the main contamination plume. The Panel notes that final guidance should include clear recommendations about sampling frequencies (see further Chapter 6).*
22	AGES (Austria)	3.2. Edge-of-field study designs for exposure assessment options 1, 2, 3, and 4	AGES does not agree that monitoring studies should not follow a rather rigid study design, particularly in the case of a targeted edge-of-field monitoring study with newly drilled monitoring wells. Usually, a harmonised study design is a key for regulatory acceptance and provides the basis for a harmonised evaluation. In this respect, AGES recommends a rather rigid study design for targeted edge-of-field monitoring studies with newly drilled monitoring wells, giving clear advice on, e.g. the minimum/maximum field size, minimum number of sampling wells, minimum number of supporting wells (assisting groundwater flow direction assessments), minimum depth to groundwater, minimum sampling frequency, minimum study duration, etc. AGES is convinced that a certain degree of 'rigidness' in the overall study design is a prerequisite for regulatory evaluation purposes and, finally, for regulatory acceptance. *PPR Panel: The Panel agrees that a 'rigid' study design would increase regulatory acceptance. A possible way forward would be developing a fixed monitoring network. This issue is address in Chapter 6 of the Statement.*
23	CropLife Europe	3.2. Edge-of-field study designs for exposure assessment options 1, 2, 3, and 4	Edge of field monitoring: (page S17) In section 'determining connectivity' the possibility is mentioned that some water in the well may not derive from the field. This section should be deleted as typically all, or the vast majority of the water will come from the field and not the 1–3 m of grass strip between the well and the field. Such a comment will add additional complexity on the connectivity discussion. In general, it would be good if the document gives some options on how to prove connectivity. At the moment some options are scattered throughout the document. For example, Connectivity at the catchment and aquifer scale can be difficult to confirm and may not always be possible. Wells actively monitored with a legacy of residue detections can provide adequate knowledge on connectivity. *PPR Panel: Arbitrary screen lengths may not reflect the groundwater able to be sampled and the screen section should be installed to capture as much groundwater variation as possible to ensure samples can be obtained. Borehole logs identify permeable material and the groundwater strike zone which should guide the screen depth and length.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
24	AGES (Austria)	3.2.1. General study outline	As indicated above, a rather rigid set-up of a monitoring location is recommended for targeted edge-of-field monitoring with newly drilled monitoring wells. E.g. it may be requested to consistently locate a sampling well at each side of a (preferably square sized) field with all these four sampling wells sampled on, e.g. bimonthly basis all the time over at least 5 years. Such a monitoring site setup minimises the risk of 'loosing' residues due to variable groundwater flow directions (being probably more the rule than the exception) and allows quantifying up-stream input to support the hydraulic connectivity assessment. Groundwater contour plots should be provided at each sampling date. In view of the overall uncertainties associated with a monitoring study as such and with contextualization assessments used to select a set of monitoring sites considered relevant for a given assessment area (e.g. a Member States). *PPR Panel: The Panel agrees that a 'rigid' study design would increase regulatory acceptance. A possible way forward would be developing a fixed monitoring network. His issue is address in Section 6.5 of the Statement*. AGES believes that the number of (fully reliable and sufficiently vulnerable) monitoring sites for a given assessment area (e.g. a FOCUS Zone or a Member State) should be as high as possible, at least 10. Assessments based on, e.g. only 1, 2 or 3 monitoring sites, as has been done in some pesticide dossiers, are considered unreliable at all. The minimum number of monitoring sites needed may be related to the extent of the assessment area. Notice, that in the case of a limited number of monitoring sites, more weight is given to individual extreme sites, and this may give unreliable assessment endpoints. *PPR Panel: See the Statement (Chapter 6) for a discussion on number of monitoring sites*. Any kind of tracers to assist the hydraulic connectivity assessment is highly welcome. Beside targeted bromide tracer experiments, it may be requested to also investigate pesticide tracers, e.g. mobile metabolites, from other pesticides applied in the up-stream area. *PPR Panel: Agreed, tracers in combination with time-of-flight modelling is the preferable method to demonstrate connectivity.*
25	UBA (Germany)	3.2.1. General study outline	Number and location of the wells. The paragraph on the determination of the groundwater flow direction reads merely like a recommendation. However, based on the experiences with targeted GW monitoring studies, the flow direction indeed might be subject to changes, especially at vulnerable sites with shallow groundwater tables. Therefore, we consider it mandatory to constantly monitor the GW flow direction (with each sampling). This is necessary because a change in the flow direction might lead to different fields being hydrologically connected to the sampling well(s) and thus questioning the reliability of the data on pesticide application at the adjacent fields. *PPR Panel: Agreed, addressed in Chapter 6*.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
26	ANSES (France)	3.2.1. General study outline	Page S17: In existing monitoring studies, limited information is usually available regarding flow direction; in addition, as indicated, flow direction can change with time. Some recommendations on how to deal with this issue would be appreciated. *PPR Panel: Agreed, addressed in Chapter 6.*
27	CropLife Europe	3.3. Catchment and aquifer designs for exposure assessment options 1, 2, 3, and 4	Catchment and aquifer monitoring: (page 18) The sub-catchment (often referred to as regional monitoring but not in this document) is squeezed and mixed in with catchment and aquifer monitoring. Given that it is an important often applied study design I feel that it deserves even its own section. Now there are several contradicting statements: For instance, it is stated that for catchment or aquifer monitoring, it would be sufficient to sample wells once. Clearly this is not the case with sub-catchment scale. No statement on the number of sites required or the sampling frequency. *PPR Panel: Comments on sampling frequency and number of sites are in Chapter 6.* Definition of the capture zone: The document should give some recommendation on it. Personally, I don't see the point of defining the full capture zone. But some estimate, also considering the length and position of the filter, should be given to demonstrate that the historic application data was collected from the right fields. Applications of tracers to all fields may not possible due to: § Permits needed § Damage to crops and soil at commercial farms § Applying at the same time as crop protection products may not be possible due to practical considerations § High background levels e.g. bromide. Therefore, tracer application to all study designs should not be made compulsory. *PPR Panel: The Panel considers that a combination of tracers or pseudo-tracers with time-of-flight modelling is the best way to demonstrate connectivity. See further Chapter 6.*
28	UBA (Germany)	3.3. Catchment and aquifer designs for exposure assessment options 1, 2, 3, and 4	We criticise that no clear recommendations are given for the site characterisation and information required for all mentioned exposure options. In some cases, necessary information cannot be provided, so that some exposure options might not be adequate to use for risk assessments. We have experiences in Germany with post registration monitoring studies comparable to option 4 and sub catchment designs with a single well per site. Such studies are conducted at 20 sites for 4–5 years on existing water quality wells when a safe use was demonstrated by modelling or lysimeter results, but uncertainties remain in Tier-1–2 risk assessments or for risk mitigation. It is required to select wells with filter screens length up to 2 m which are installed 0–2 m below the GW table in areas with a GW table > 1 and < 5 m below surface. Drilling profiles, filter screen positions, water contour maps and geological information are provided by the authorities responsible. PPP use information is collected from farmers. Weather and soil data are sometimes additionally collected. Tracer experiments are usually not conducted. We must admit that PPP residues in shallow groundwater can be detected with this study design but demonstrating connectivity during study time is mostly difficult. The one-well-design hampers to

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			derive the exact groundwater flow direction over time. Precise hydrological site information is often not available to reliably estimate leaching times in the unsaturated zone and transport times in the saturated zone. The exclusion of false negatives is not always possible. We therefore recognise that high uncertainties remain for (sub)catchment, aquifer designs with single wells to demonstrate connectivity for specific PPP uses. The edge of field design with more than one well per site is much more appropriate to overcome those uncertainties when monitoring results are quantitatively used for highest tier groundwater risk assessments according to intended PPP uses. *PPR Panel: Recommendations on study design are given in Chapter 6*.
29	ANSES (France)	3.3. Catchment and aquifer designs for exposure assessment options 1, 2, 3, and 4	The determination of the contributing catchment area is one of the key factors when assessing a monitoring study. It would be helpful to provide some recommendations regarding options for determining the contributing catchment area. *PPR Panel: We assume that you are talking about the connectivity issue here. We give some recommendations on this issue in Chapter 6 of the Statement. However, final guidance development can only be done after the ExAG has been determined by risk managers*.
30	UBA (Germany)	3.4. Study designs for exposure assessment option 5	The authors confirm that according to the exposure assessment option 5 (in table 1) where the concentration limit applies to groundwater deeper that 10 m below the surface, two different types of study sites would be possible to select: First, sites with a water table close to 10 m below the surface and second, sites with a water table quite shallow (e.g. 1–2 m below the ground surface). This is irritating and not very precise, because both potential sites mentioned here have very different hydrological characteristics and the behaviour of PPPs in the saturated and the unsaturated zones would be different at both sites. Answering the questions to define a protection goal (see Appendix 1), both example sites would probably end up in different exposure assessment options. Therefore, and as already stated in our comment on chapter 2 (table 1), we would like to recommend again that defining an exposure assessment goal in future always requires to define an acceptable depth of the groundwater table (or a range because it can change over time) in relation to the ground surface in addition to the definition of the filter screen depth. The filter screen depth should always be defined in relation to the groundwater table and not only in relation to the ground surface. *PPR Panel: We agree that depth is an important part of the Exposure Assessment Goal, considerations are given in Chapter 2 of the Statement. Our most important concern is that it should be protective for the given exposure assessment goal. The spatial unit of the Exposure Assessment Goal is to be defined by the depth of the groundwater unit below soil surface as well as by the depth interval in the upper groundwater layer of the groundwater unit*. C11: Because of expected longer times for PPP leaching in the unsaturated zone and PPP transport in the saturated zone it will be rather difficult with the mentioned study designs to reliably demonstrate connectivity and to exclude false negatives (see our comment to chapter 3.3). We therefore think that the recommended study

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			designs for option 5 are rather not suitable to collect monitoring data of sufficient quality to be quantitatively used for highest tier groundwater risk assessments according to specific intended uses of PPPs. *PPR Panel: We tend to agree that a monitoring depth between 1 and 10 m is more adequate because of the issues mentioned. This does not mean that risk managers can set the exposure assessment goal at 10 m depth. Monitoring can also be designed to give a worst-case estimate of the given protection goal. Considerations on connectivity are given in Chapter 6.*
31	UBA (Germany)	3.5. Study designs for exposure assessment options 6 and 7	Because of expected long times for PPP leaching in the unsaturated zone and PPP transports in the deeper saturated zone it will be hardly possible to demonstrate connectivity for exposure assessment options 6 and 7 in an acceptable time frame, because drinking water pumping stations are often much deeper installed than 10 m below the surface or even below the groundwater table (for example in Germany). Any detections of PPP in those wells will therefore be integrated values from large catchments and from PPP applications several years or decades before. Excluding false negatives cannot be ensured at all. We acknowledge that monitoring studies according to option 6 and 7 could be interesting for risk management purposes in areas with findings of high concentrations. However, we think that exposure option 6 and 7 do not qualify for groundwater risk assessments according to Regulation (EC) 1107/2009. *PPR Panel: See answer to comment #30.*
32	ANSES (France)	3.5. Study designs for exposure assessment options 6 and 7	Page S22: The following is reported: 'since several years is often required for water to move to the inlet of a drinking-water pumping station, rapid changes in the groundwater concentrations should not occur. Therefore, samples collected at a single time should be adequate to demonstrate lack of an active substance or its metabolites'. This could be clarified, since sampling collection at a single time does not seem consistent with appendix I options 6 and 7 in which daily/weekly concentrations are mentioned. *PPR Panel: This is not further discussed because we see ExAG 6 and 7 as not relevant in the framework of Regulation 1107/2009.*
33	CropLife Europe	4. Groundwater vulnerability assessment and mapping	p22 Groundwater vulnerability varies across Europe and may be different in different Member States. Specific groundwater protection is more appropriate at a Member State level, where different protection goals may exist. An equivalent FOCUS pass and therefore 'one safe use' might be more appropriate at an Annex I level. *PPR Panel: We don't see why an exposure assessment goal would not be necessary at the EU-level. Also, for determining a safe use, an exposure assessment goal is needed. Furthermore, we would recommend harmonising protection among Member States; however, this is a political decision.*
34	CropLife Europe	4.1. Groundwater vulnerability concepts	Intrinsic vulnerability is a theoretical concept which may be better defined as the physical variables which can influence the rate of leaching of an inert and chemically unreactive substance, rather than all solutes as stated in this report. Not every solute present in the system will respond to the physical variables present in the same way e.g. the response to soil organic carbon will be dependent upon the sorption properties of a compound. There are a range of techniques already established to assess vulnerability and their suitability in Europe should be

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			examined. Aquifer characteristics should be assessed at MS level taking into consideration the Tier-4 study design. Shallow groundwater studies often reflect the most vulnerable groundwater. *PPR Panel: The assessment of groundwater vulnerability must consider substance properties and application patterns. So intrinsic vulnerability is not enough. See Chapter 4.*
35	CropLife Europe	4.1.1. Intrinsic/ environmental vulnerability	'For some crops where the natural precipitation is insufficient or too irregular, artificial irrigation is used. This should be considered similarly to precipitation in the context of the contribution to leaching potential'. Spatial information on appropriate irrigation amounts may be difficult to obtain. *PPR Panel: We agree that irrigation must be considered in the in-context setting procedure. See Chapter 4.*
36	CropLife Europe	4.1.2. Specific vulnerability	'However, substance parameters (DT50, Koc) will also play a role as they will determine the timeframe in which an active substance or metabolite will be present in the soil, and at which depth, following an application'. It can be assumed that hydrolysis will occur throughout the soil profile (unsaturated zone) before the leachate will reach the groundwater level/aquifer. Should this additional degradation pathway of the active substance or metabolites not be considered as well? *PPR Panel: All processes that are considered in Tier-2 can be taken on board in higher tiers as well. See Chapter 4.*
37	UBA (Germany)	4.1.2. Specific vulnerability	Groundwater monitoring studies performed in the past often tend to reflect so called realistic use conditions. Information on the dose rates, the intervals between applications, the hydraulic connection between the monitoring well and the adjacent fields and the stability of the groundwater flow direction is often limited. However, for any regulatory assessment in terms of risk, groundwater monitoring results need to be directly related to the representative use they are intended to cover (according to the GAP). This crucial issue is mentioned at different chapters of the paper. We are aware that the paper summarises the scientific state of the art, but from a regulatory perspective the information provided lack precision and consistency. In our view, the following information is essential in the regulatory assessment of any targeted GW monitoring study: – application rate (in g/ha), – application pattern (e.g. biannual, annual, biennial, triennial...), – proof of hydraulic connection between the monitoring well and the adjacent fields, – steadiness of the groundwater flow direction. The data on application (rates, pattern) is needed to relate the monitoring data to a pre-defined representative use. The evaluation of both the application rates and the patterns is still a work in progress, further discussions and guidance are urgently needed. As under practical agricultural conditions the strict boundary conditions as defined in a representative use are seldomly fulfilled in a 1:1 manner, approaches need to be developed how to decide upon the question which application rates and patterns the results of the monitoring do cover. The data on connectivity and groundwater flow direction over time is required to provide certainty that the available data on application is indeed relevant for the respective monitoring well over the course of the whole study duration. Only if this information issufficient the GW monitoring results may be used for regulatory risk assessment

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: We agree that any future groundwater monitoring studies should include application rates, application frequency, application patterns, etc. Also scaling of historic uses to the GAP should be done in a harmonised way. We address these issues in Chapter 6.
38	ANSES (France)	4.2. Vulnerability mapping approaches	Vulnerability mapping approaches are specific to one compound. However, in many cases, monitoring studies are performed for several compounds (parent substance and several metabolites), which do not necessarily have the same properties (degradation, mobility). Sites selection when several compounds need to be monitored should be further discussed. An analysis of the vulnerability can be done for each metabolite, and it may occur that the different compounds cannot be handled by an identical site. That could highly increase the number of sites. *PPR Panel: We recommend selecting the monitoring site based on the substance of most concern. In context setting for the other substances can be done afterwards (Chapter 4).* As indicated, vulnerability mapping approaches are usually limited to the unsaturated zone. Although it is possible, coupling with models for saturated zone remains a complex approach and will be difficult to consider in the framework of Regulation 1107/2009 assessments. *PPR Panel: We recommend that the vulnerability assessment is limited to the upper meter. Other items (e.g. connectivity) must be dealt when interpreting the monitoring study (Chapter 6).* The use of mass fluxes rather concentrations in vulnerability mapping is proposed in the example presented. The rationale behind this choice could be discussed in further details and clear recommendations for further application could be provided. *PPR Panel: This issue is addressed in Chapter 4.*
39	CropLife Europe	4.2.1. Scope of vulnerability assessment and mapping	'At what spatial scale is the vulnerability to be assessed?' Vulnerability modelling should be consistent and across the same spatial scale to ensure no bias. *PPR Panel: The spatial scale must be uniform to avoid bias (Chapter 4).*
40	UBA (Germany)	4.2.1. Scope of vulnerability assessment and mapping	It is not clear from the introduction in this chapter what is meant by 'the temporal scale of the vulnerability' and how this could be considered by choosing methods and appropriate data to assess or map vulnerability. *PPR Panel: It's also not clear to us either.*
41	AGES (Austria)	4.2.3. Different vulnerability mapping approaches	Process-based vulnerability mapping requires a leaching model, such as one of the FOCUS models, to estimate the groundwater leaching vulnerability at local level. This leads to the fundamental question whether a process-based leaching model with an assessment depth of 1 m soil depth is adequate to describe the real-world groundwater leaching vulnerability at local level. AGES has no strong objections on using a spatially distributed FOCUS model to derive an 'anonymous' leaching distribution for a certain assessment area (e.g. a Member State) and to calculate a certain spatial percentile concentration in sense of FOCUS Tier-3b (e.g. at 1 m soil depth). However, in view of AGES, there is a high risk to fail the real-world leaching to groundwater (and to monitoring

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			wells) for individual locations at this distribution curve, because local substance properties may be different from generic 'average' properties and/or unknown processes below the assessment depth are not accounted for. In this respect, AGES does not necessarily support the use of a process-based model to 'quantify' and to compare the leaching vulnerability of individual monitoring sites or to select certain monitoring sites as being highly vulnerable, e.g. for a Member State. In this respect, an index-based approach with a focus on the intrinsic/ environmental (i.e. pedoclimatic) vulnerability may be less controversial and probably more in line with regulatory needs. Figure 20 in Gimsing et al. (2019) may give an indication on how uncertain a leaching vulnerability mapping based on a process-based model is: In this case, there is a factor of only 2.5 in the calculated mass flux between the most vulnerable and the least vulnerable monitoring locations all over Europe. AGES has strong doubts whether such 'flat' vulnerability maps have any meaning for real-world groundwater vulnerability at all and at given monitoring sites. *PPR Panel: We don't agree that index models could replace process-based models in vulnerability assessment, because substance properties play an important role in groundwater vulnerability. For the in context setting procedure and the site selection procedure it doesn't matter that substance properties are variable. They average out when looking at a spatial aggregate. Chapter 4.*
42	CropLife Europe	4.2.3. Different vulnerability mapping approaches	Vulnerability mapping is a convenient means by which potential groundwater residues in one area can be compared to another and can be useful for selecting monitoring sites. It is not clear how the results of these studies could be used to support the 'one safe use' concept used to make decisions on groundwater risk at a European level. One option would be to use these data to demonstrate an 'equivalent FOCUS pass' as outlined by Sweeney (22nd Fresenius conference 2020). *PPR Panel: This issue is addressed in Chapter 4 (in context setting).*
43	UBA (Germany)	4.2.3. Different vulnerability mapping approaches	C15: The authors compare index- & process-based models to be used for vulnerability analysis und conclude that 'process-based models are a convenient way of integrating the environmental factors that affect leaching and quantifying potential residues, so that the relative leaching vulnerability of one location can be compared to another'. GeoPEARL, EuroPEARL, MACRO-SE, Proziris are mentioned as process-based models 'which can be used for generating vulnerability maps'. (p. S26). We agree that degradation and sorption of active substances and metabolites in soil are probably addressed in a better way compared to index-based models. However, considering the fact mentioned in chapter 4.1.1 (p. S24f) that preferential flow can play a major role for leaching, especially in fine-grained soils, the different models listed here will produce very different outputs calculating the spatially distributed vulnerability. While the MACRO-based models account for preferential flow, the PEARL-based models do not. It can be expected that a vulnerability map calculated with PEARL will rank coarse-grained (sandy) soils with a higher vulnerability than coarse-grained (silty, clayey) soils. The outcome will be different and maybe opposite from MACRO-based models where both fine- and coarse-grained soils are expected to be vulnerable. Consequently, the spatial cumulative distribution functions and vulnerable areas identified will strongly vary among both model types. We identify this as a contradiction in the recommendations which needs to be

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			solved in further guidance development. We regard preferential flow an important process to be considered if process-based models are used for vulnerability estimations. A research project investigating spatial distributed modelling with PELMO in Germany (report in prep.) provides evidence that any other process implemented in leaching models, especially preferential flow, will strongly influence results of vulnerability mappings and rankings. *PPR Panel: This issue of preferential flow is discussed in Chapter 3 and 4.*
44	CropLife Europe	4.2.5. Spatial data considerations	Geodata which is appropriate for spatially distributed leaching modelling (SDLM) is already listed in the problem definition document of the SETAC SDLM group. Furthermore, geodata is currently under evaluation of the SETAC SDLM geodata subgroup. The SETAC SDLM working group will come up with a recommendation on which current data sets should be used for SDLM and how the data can be schematized to reduce runtime. SETAC SDLM will also recommend measures to ensure the quality of the geodata including version control for geodata as it is the case for FOCUS models. *PPR Panel: Noted.*
45	CropLife Europe	4.2.6. Technical considerations	It is important that the vulnerability is not a function of the data itself if comparison between one area and another is to be meaningful. If different environmental data are used that cover different spatial scales, or have been derived using different methodologies then modelled vulnerability might be a function of these data rather than the intrinsic properties of a molecule or use *PPR Panel: Agreed and therefore harmonisation of geodata in vulnerability assessments is needed (see example for the EFSA soil exposure guidance (2014) where the EFSA dataset is used). See further Chapter 4.*
46	AGES (Austria)	4.2.7. Geoinformation sources	In view of AGES, there is an urgent need to define a harmonised dataset on geoinformation which must be used by each applicant and each evaluator. There is also a need for a harmonised FOCUS zone map. *PPR Panel: Agreed and therefore harmonisation of geodata in vulnerability assessments is needed (see example for the EFSA soil exposure guidance (2014) where the EFSA dataset is used). See further Chapter 4.*
47	CropLife Europe	4.2.7. Geoinformation sources	Geodata which is appropriate for spatially distributed leaching modelling (SDLM) is already listed in the problem definition document of the SETAC SDLM group. Furthermore, geodata is currently under evaluation of the SETAC SDLM geodata subgroup. The SETAC SDLM working group will come up with a recommendation on which current data sets should be used for SDLM and how the data can be schematized to reduce runtime. SETAC SDLM will also recommend measures to ensure the quality of the geodata including version control for geodata as it is the case for FOCUS models. *PPR Panel: Noted.*
48	AGES (Austria)	4.3.1. Monitoring site characterisation and vulnerability assessment	Identifying false positive, false negative and true negative monitoring results may be extremely challenging, particularly for evaluators having limited access to data. Any further guidance addressing this issue with clear criteria is highly welcome.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: Agreed that that guidance is needed for the interpretation of these studies. This includes of course the identification of false positives/negatives etc. See further Chapter 6 and 8.
49	UBA (Germany)	4.3.1. Monitoring site characterisation and vulnerability assessment	Row on 'Use intensity' in table 2 on page 33 is too unspecific. Sales data are not sufficient for the assessment of (targeted) GW monitoring results. Detailed data on the pressure of use are required (application rates and patterns). Please refer also to the comment on chapter 4.1.2. (Specific vulnerability: pressure of use). *PPR Panel: We agree that any future groundwater monitoring studies should include application rates, application frequency, application patterns, etc. Also scaling of historic uses to the GAP should be done in a harmonised way. We address these issues in Chapter 6.* It is mentioned in this chapter (and in a more general way already in chapter 4.1.1: saturated zone, p. S24) that monitoring site characterisation and vulnerability assessment depends on hydrogeological parameters (listed on p. S32). We think that hydrogeological conditions can be very divers across Europe and that they can influence GW monitoring results to a high extent. They should therefore be considered in vulnerability assessments with a comparable weighting like other intrinsic properties, e.g. soil and climate conditions. However, to our knowledge, the European-wide current availability of hydrogeological parameters is very limited, and even the availability on a national scale can be difficult. To give an example, nation-wide data about the depth and fluctuation of the GW tables are even not available in Germany in a scale which would be necessary to identify all potential vulnerable areas in Germany. We recommend that the availability of hydrogeological data should be more critically discussed regarding solutions for suitable vulnerability assessments in future. *PPR Panel: We agree that subsoil properties may play an important role in assessing the vulnerability of an aquifer. However, given the data availability, the Panel recommends limiting the vulnerability assessment to the top 1-m. An important justification for this is that most dissipation processes take place in this layer (Chapter 4). However, we recommend that after a site has been selected, the geohydrological situation must be investigated and considered. Further guidance development on this issue is required (Chapter 6).*
50	ANSES (France)	4.3.1. Monitoring site characterisation and vulnerability assessment	Page S31: in the 4th paragraph (right column), it is written 'If the exposure assessment option considers deeper groundwater (as in options 4–7) ...'. However, option 4 does not include sampling in 'deeper groundwater' but rather sampling in shallow groundwater, with sampling depth between 1 and 10 m. It is therefore assumed that '(as in options 5–7)' should have been indicated. *PPR Panel: Agreed, this appears to be an error in the article.*
51	AGES (Austria)	4.3.2. Assessing the hydraulic connectivity between sampling points and treated fields	In view of AGES, adequate hydraulic connectivity is a key in considering sample results from individual monitoring wells reliable or not. Due to variable groundwater flow directions, nominated sampling wells may be temporarily or even permanently unconnected to a treated field, particularly at monitoring sites where multiple sampling wells have been installed to address variable groundwater flows directions. Therefore, AGES highly recommends developing more targeted guidance on how to demonstrate and how to quantify adequate connectivity between treated field(s) and sampling well(s), e.g. by means of groundwater contour plots, transducer data, hydraulic

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			conductivity data, tracer experiments, etc. There is a need for clear criteria on when to consider a sampling well adequately connected and when not, and how to proceed if a sampling well is not adequately connected. Such criteria may, e.g. account for a minimum percentage of samples required with the sampling well adequately located in the groundwater flow direction (e.g. based on groundwater contour plots obtained at each sampling date). Considering the key position of adequate hydraulic connectivity in the overall evaluation of a monitoring study, neither a 'weight-of-evidence' approach nor 'high-tech' modelling (e.g. 2D- or 3D-modelling, etc.) is considered appropriate in a regulatory framework. *PPR Panel: Agreed that connectivity needs attention in guidance development. This is addressed in Chapter 6.*
52	CropLife Europe	4.3.2. Assessing the hydraulic connectivity between sampling points and treated fields	When instrumenting new sites, groundwater flow direction needs to be understood via triangulation rather than relying on topography or potentiometric maps. Other long term established well networks in MS may not need to ascertain groundwater flow in the same was as a newly installed site, due to the legacy of data and already known understanding of groundwater flow. Reference/datum points should be recorded for several permanent features of a well e.g. casing, piezometer and ground level. At commercial fields, well location may not be ideal to capture all groundwater flow and locations can be compromised to satisfy landowner's needs. *PPR Panel: Guidance development on connectivity is needed. This is addressed in Chapter 6.*
53	UBA (Germany)	4.3.2. Assessing the hydraulic connectivity between sampling points and treated fields	The recommended method 'time-of-flight' to estimate connectivity for a certain time frame requires intensive data input and complex site-specific modelling. We acknowledge the scientific recommendations as possible way to overcome missing hydrogeological knowledge and/or tracer experiences at newly selected sites/wells. However, we have concerns that this modelling procedure adds undesirable complexity for future assessments of targeted monitoring studies. Besides, uncertainties in the estimated connectivity will remain when the used input data for modelling are not precisely known. A stable and fixed targeted monitoring system in Europe with fixed and suitably characterised wells (e.g. by tracer experiments) could reduce the effort and complexity in future GW monitoring studies in that regard. (See our comment C3 in chapter 3) *PPR Panel: Agreed that connectivity needs attention in guidance development on groundwater monitoring. The issue of fixed monitoring sites is addressed in Chapter 6.*
54	AGES (Austria)	4.3.3. Application of vulnerability mapping for the identification of potential monitoring sites	AGES does not see a need for further guidance (or harmonised model approaches) on how to identify potential monitoring sites. Processed-based model may support such an approach, but there is of course no obligation to apply highly sophisticated models in the site selection procedure, other approaches, e.g. index-based ones, may be similar successful. Finally, it always must be demonstrated that monitoring wells (newly drilled or already installed) are adequately connected to the treated field(s), are sufficiently vulnerable for the specific assessment goal and have application rates and frequencies close or equal to the intended one. *PPR Panel: We don't agree that index methods could replace process-based models or statistical metamodels because these don't consider the dependence of vulnerability on substance properties. See Chapter 4. The connectivity issue is described in Chapter 6.*

efsa JOURNAL

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
55	CropLife Europe	4.3.3. Application of vulnerability mapping for the identification of potential monitoring sites	In the context of vulnerability mapping for the identification of potential monitoring sites mass flux is mentioned as appropriate metric rather than flux concentrations at a certain soil depth. This is in agreement with considerations of the SETAC SDLM working group in the problem definition document. This topic will be further discussed in the SETAC SDLM working group. *PPR Panel: This issue is discussed in Chapter 4.*
56	UBA (Germany)	4.3.3. Application of vulnerability mapping for the identification of potential monitoring sites	The recommended application of vulnerability mapping to identify potential monitoring sites, explained in figures 16–19 by an example calculated with GeoPEARL, does not consider the influence of preferential (or macropore) flow for vulnerability mapping (see our comment C15 in chapter 4.2.3). Different vulnerable areas could be identified if preferential flow would be included in such modelling routines. We think that vulnerability mapping needs more scientific development before it can be used for a site selection across Europe. A validation of different modelling approaches with groundwater monitoring data should be aspired for further developments of vulnerability mapping. PPR Panel: The issue of preferential flow is discussed in Chapter 2 and 4. The example by Syngenta shows the current state of the art in terms of vulnerability mapping. Nevertheless, several uncertainties are introduced using models and heterogeneous input data across the EU. For example, the groundwater depth data (quality and availability vary across the EU) used for the vulnerability assessment need to focus on the most upper aquifer, regarding the chosen exposure option. The example by Syngenta (2014) only specifies that depth to groundwater should be < 10 m. *PPR Panel: Depth is an important part of the exposure assessment goal and should be defined by risk managers. See Chapter 2.*
57	ANSES (France)	4.3.3. Application of vulnerability mapping for the identification of potential monitoring sites	The context setting of monitoring sites is an interesting approach. However, this part may need further analysis beyond the description of the methodology developed by applicants for vulnerability context setting. For example, figure 20 comparing the median flux calculated for the monitoring locations (based on actual site characteristics) vs the median flux calculated for the whole EU zone should be further analysed. It can be seen on this plot that, while the selected sites well cover the whole range of leaching vulnerability situations that is encountered throughout Europe, many sites are below the top 60th percentile of median annual flux, which was the cut-off value used to build the vulnerability mapping for site selection. It implies that many of the sites selected have a clear lower level of leaching vulnerability than the targeted level, and the overall vulnerability represented by all monitoring sites is far below a 60th percentile. Criteria regarding the suitability of the distribution could be proposed. This should be considered further in analysis of monitoring results, and criteria on percentile of vulnerability covered by the actual selected sites needs to be included in the development of the options. *PPR Panel: We agree that guidance on the in context setting procedure is needed. See further Chapter 4.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
58	AGES (Austria)	4.3.4. Application of vulnerability mapping for setting monitoring sites into context	See comments given at 4.2.3. *PPR Panel: Noted.*
59	UBA (Germany)	4.3.4. Application of vulnerability mapping for setting monitoring sites into context	The authors conclude that vulnerability modelling 'allows the direct comparison of the pedoclimatic leaching vulnerabilities of the monitoring sites in the context of the area of interest. (...) In this way, the monitoring data can be used to draw conclusions about the probable leaching risk to groundwater in areas where no monitoring data are available'. (p. S37) We refuse vulnerability modelling in the way it is recommended (e.g. in figure 20) as too imprecise to set specific monitoring sites into context of large areas of interest. This is for the following reasons, which belongs also to chapter 4.4: The modelling routines in the recommended leaching models do not apply for all processes influencing the water balance in the unsaturated soil zone, e.g. preferential flow and runoff, which frequently occur in the fields and can influence the monitoring results. Climate data in the recommended models refers to previous decades. It cannot be assumed under the dynamic of climate change that actual weather conditions during a monitoring study are easily comparable to the historic climate data used in the models. The soil parametrisation in the recommended models is normally derived from generic soil maps which do not represent a field scale. Soil characteristics from monitoring fields can vary and are not automatically comparable. The recommended vulnerability approach considers no aquifer conditions from the saturated zone which additionally influence the monitoring results. A large variety of field characteristics influence monitoring results, whereas models are simplified approaches which cannot account for all parameters. Results from spatial vulnerability ranking by modelling are less accurate to interpret real monitoring results as suggested by the authors. We have concerns that this method will be used to demonstrate vulnerability of only few monitoring sites instead of defining a sufficient extended site/well selection across Europe. *PPR Panel: We argue that there is currently no true alternative to modelling for in context setting of monitoring sites. The currently used models have successfully been applied in validation studies across Europe and we consider them therefore fit for purpose. We do agree, however, that the models also have limitations, particularly when simulating preferential flow. There is, however, discussion if this is really a problem in the vulnerability assessment. E.g. sites with clay soils are often artificially drained and a high fraction of the mass transported by preferential flow may not reach the groundwater. On the contrary, sandy soils are less frequently drained. So, we agree that further research on the effect of preferential flow on vulnerability assessment is needed (see Chapter 4). We further agree that field data may be different from generic data, based on soil maps. We therefore suggest splitting the site selection into two phases, i.e. a preselection phase based on modelling with generic parameters and then the final selection in which the applicant must demonstrate that the site conditions match the generic conditions. The problem of climate-change should be overcome by using recent climate data. The Panel therefore suggests regular updates and version control.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
60	AGES (Austria)	4.4. Interpretation of spatial vulnerability assessments and context setting of monitoring sites	AGES acknowledges that in this section uncertainties associated with vulnerability maps and context setting of monitoring sites are discussed more critically (also refer to comment given at 4.2.3). As indicated above, AGES has concerns about vulnerability mapping and subsequent context setting if there is no consensus that this approach is sufficiently robust for regulatory purposes. In view of AGES, uncertainties arising from such approaches should be evaluation more in detail on basis of monitoring data available for numerous pesticides. AGES wants to make aware that at the level of Member States a 90th spatial/temporal percentile concentration, obtained from a groundwater monitoring study, may not be acceptable, depending on the groundwater protection goal in a Member State. Groundwater is an entirely different compartment compared to pore water at 1 m soil depth considered at FOCUS Tier-1 and there is no 'safety factor' or conservatism left at FOCUS Tier-4. In this sense, a 90th spatial/temporal percentile concentration in groundwater at FOCUS Tier-4 may have a completely different meaning and different regulatory consequences compared to a 90th spatial/temporal percentile concentration in soil pore water at FOCUS Tier-1. *PPR Panel: The selection of the spatiotemporal percentile is a risk management decision because it also includes socio-economic considerations. However, selecting a percentile above the 90th-spatiotemporal percentile may challenge the tiered approach. This issue is discussed in Chapter 2 (protection goal) and Chapter 3 (tiered approach).*
61	CropLife Europe	4.4. Interpretation of spatial vulnerability assessments and context setting of monitoring sites	This is a key concept for using groundwater monitoring studies for higher-tier risk assessment. Following the underlying concept of the modelling scenarios at lower tiers, the monitoring locations can be considered to represent a pedoclimatic leaching risk. Using spatial data/modelling to determine the representativity of the monitoring locations for a given area of interest (e.g. country or FOCUS climate zone) follows essentially the same logic as the derivation of the FOCUS groundwater scenarios and is compatible with the FOCUS risk assessment concept. Important to note is, that a given monitoring location can be representative for different areas of interest. '... a reasonable approach is to consider the 80th temporal percentile of measured concentrations from locations corresponding to the 80th percentile pedoclimatic leaching vulnerability for the area of interest'.To be able to compare the results of Tier-1 to Tier-3 (modelling) with Tier-4 results (monitoring) it is required to follow the same evaluation approach. A monitoring study conducted on fields above shallow groundwater a few metres deep should address concerns over deeper aquifers and aquifer characteristics as these are the most vulnerable to chromatographic flow. *PPR Panel: Comment is generally agreed (see also Chapter 3 of the Statement). However, the selection of the spatiotemporal percentile is a risk managers decision (Chapter 2).*
62	ANSES (France)	4.4. Interpretation of spatial vulnerability assessments and context setting of monitoring sites	Page S39: When vulnerability mapping approaches are provided, it should be recommended that an analysis of the uncertainties related to the approach used is systematically submitted. *PPR Panel: Agreed and guidance for addressing this uncertainty must be developed. Example from the Good Modelling Practice-opinion (EFSA PPR Panel, 2014) could be a starting point.*

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
63	UBA (Germany)	4.5. Generic recommendations for vulnerability assessment and site characterisation in monitoring study design and interpretation	In our view chapter 5 is well elaborated and provides a variety of scientifically sound options for many purposes and addresses potential problems and constrains with pragmatic and reliable solutions. However, regarding a guidance document a clear definition and binding procedures must be defined to produce adequate studies for risk assessment, e.g. in which cases are available information sufficient and in which are tracer tests or field spikes necessary (please refer also our other comments on this). Particularly in chapter 5 the assignment of requirements, quality criteria and procedures to the respective exposure option (protection goal) and study design is not sufficiently provided. For instance, for deeper groundwater (exposure options 5, 6, 7) and catchment study designs a profound characterisation of (hydro-)geology is essential, while in shallow groundwater and in field studies this might be of minor relevance. In the latter, contamination during study performance or local/temporal land use changes might be more important. Regarding a guidance, these requirements must be defined. *PPR Panel: Comment doubles with comment #64.*
64	UBA (Germany)	5. Data quality considerations	In our view chapter 5 is well elaborated and provides a variety of scientifically sound options for many purposes and addresses potential problems and constrains with pragmatic and reliable solutions. However, regarding a guidance document a clear definition and binding procedures must be defined to produce adequate studies for risk assessment, e.g. in which cases are available information sufficient and in which are tracer tests or field spikes necessary (please refer also our other comments on this). Particularly in chapter 5 the assignment of requirements, quality criteria and procedures to the respective exposure option (protection goal) and study design is not sufficiently provided. For instance, for deeper groundwater (exposure options 5, 6, 7) and catchment study designs a profound characterisation of (hydro-)geology is essential, while in shallow groundwater and in field studies this might be of minor relevance. In the latter, contamination during study performance or local/temporal land use changes might be more important. Regarding a guidance, these requirements must be defined. *PPR Panel: Agreed that guidance development is needed. Addressed in Chapters 5 and 6 of the Statement.*
65	UBA (Germany)	5.6. Collection of samples	The paper states that 'sampling procedures for groundwater studies conducted in a specific situation or directed towards a specific objective, may not be suitable for studies in other situations and with different objectives' (p. S44). We agree on that; however, a subsequent guidance document should clearly define which sampling procedure is recommended in which situation regarding which protection goal and study design (please refer to our general comment on chapter 5). *PPR Panel: Agreed that guidance development is needed. Addressed in Chapters 5 and 6 of the Statement.*
66	UBA (Germany)	5.6.3. Water removal	On the particulate phase of groundwater samples, the paper states: 'While including solids in samples should always be avoided, this is especially important for compounds strongly sorbed to solids (active substances and metabolites with high Koc values). Even though they rarely move through soil in groundwater, sometimes analyses of strongly sorbing materials are included and it is important to exclude solids to get reliable analyses for

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			these compounds' (p. S45). From a water analysis perspective, solids shall of course be removed (filtrated) before analysing the samples. However, strongly adsorbing compounds are mentioned. Bound to particles those compounds might be transported to groundwater via macropore flow. Could it be reasonable to also analyse the solid phase (filtrate) on strongly adsorbing compounds if these compounds are objective of the study? A subsequent guidance document might define for which substances (e.g. K_{oc} above a certain value) an analysis of the solid phase will be reasonable and recommended. *PPR Panel: Sampling solids should be avoided because we are interested in the aquatic concentration. However, if we are interested in colloid-facilitated transport, it could be an option to sample the filtrate. See Chapter 5.*
67	UBA (Germany)	5.6.7. Sampling other types of wells	On sampling of other well types, the paper states 'for example, sampling a well providing water to a private residence is often as simple as turning on the faucet, letting it run for a specified period of time (e.g., 30 s or 1 min), triple-rinsing the bottle and filling the sample bottle' (p. S46). This sampling technique is of course possible to get an impression on groundwater quality if there is no other sampling possibility. We agree with the authors that there are many uncertainties and potential sources of contamination. Regarding a guidance document, we would conclude that studies based on this kind of sampling cannot be used as Tier-4 studies for risk assessment. *PPR Panel: Agreed, this simple sampling technique should not be used in a Tier-4 assessment. This is addressed in Chapter 5.*
68	CropLife Europe	5.7. Sample analysis	The lab guidelines are stated even with revision numbers. May be only state the guidelines as these will change over time. SANTE is not mentioned but should be applicable for method validation. *PPR Panel: Noted.*
69	UBA (Germany)	5.8 Outliers	Groundwater monitoring studies intend to measure the groundwater concentrations caused by leaching processes. Therefore, it appears plausible to exclude results that are driven by other factors than leaching through the soil column following product uses according to the good agricultural practice. Chapter 5.8 explicitly recommends additional detailed investigations of sites with 'unexpectedly high' detections and lists a few possible reasons for false positives. However, we have serious concerns to investigate only the sites with extraordinarily high concentrations. To exclude those sites from datasets may introduce a bias to the overall picture, because sites with no or very low detects are not treated with the same attention. There may be similar reasons that can result in no/low-detects (false negatives) which are not related to leaching, e.g. – dry weather conditions, – non-agricultural upstream areas, – lack of connectivity to treated fields during the sampling period, – change in groundwater flow direction, – incorrect data on product use, – clogged filters, – storage and stability issues, – analytical problems. Therefore, we recommend treating 'unexpectedly high' and low detects with the same caution. We would like to encourage further discussions and appreciate the development of further detailed guidance on handling false negatives and false positives. This comment refers also paragraphs on the issue in chapters 4.3.1 and 7.1.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: Agreed that both false positives, and false negatives should be carefully investigated. Some examples of both will be included in Chapter 6. Guidance needs to be developed.
70	CropLife Europe	5.9.1. Tracers	Applications of tracers to all fields may not possible due to: Permits needed, damage to crops and soil at commercial farms, applying at the same time as crop protection products may not be possible due to practical considerations, high background levels of tracer, e.g. bromide Therefore, tracer application to all study designs should not be made compulsory. *PPR Panel: We see the application of tracers or pseudo-tracers as the preferred methodology. If tracers cannot be applied, modelling is an alternative, but the Panel recommends strict guidance development to support this (Chapter 6)*.
71	ANSES (France)	6.1. Assessing groundwater monitoring data	Page S50: Data from surveys on agricultural practices often show that the application pattern (rate and frequency) on the monitored fields included in the contribution area is lower than the intended application pattern of the formulation evaluated (for which maximum dose is assessed). In addition, in some cases, the information on application is not available for the whole surface of the defined contribution area. Recommendations on how to deal with these issues should be provided. *PPR Panel: For future studies, the Panel recommends these data to be made available. If historical application rates differ from the application rates according to the GAP, scaling is possible. This is described in Chapter 6*.
72	INIA (Spain)	6.1.1. General considerations	The article provides a great review of specific exposure assessment options to be considered at higher-tier groundwater risk assessment (Tier-4). Recommendations for study designs, quality criteria of monitoring studies and assessment and interpretation of groundwater vulnerability have been detailed. We consider that it is an important step in the assessment of monitoring studies. However, it is not clear when monitoring data can be used to override the results from standard FOCUS groundwater models. We highlight from the report that there is an urgent need of defining and agreed at EU level specific protection goals to groundwater monitoring data. *PPR Panel: Groundwater monitoring studies can override lower tiers if they have been well performed. Some recommendations for this are given in Chapter 6 but final guidance needs to be developed in a later stage. We agree that exposure assessment goals must be defined before guidance can be developed (Chapter 2)*.
73	CropLife Europe	6.1.1. General considerations	It should be noted that exceedances are permitted in Tier-1 modelling, where, in principle, 3 out of 20 exceedances could be permitted. The FOCUS report recommendation for 20 targeted sites to demonstrate safe use at Tier-4 does not have a formal statistical basis but is certainly a reasonable number and practically manageable. If a leaching risk is not seen at 20 carefully selected vulnerable and well characterised locations in a monitoring study, then it is unlikely to be seen if further monitoring sites are added. From a practical point of view, finding suitable monitoring locations that fulfil all requirements of quality and intrinsic vulnerability while at the same time having relevant crop cultivation and product usage, and where the landowners are cooperative, proves to be very challenging. The effort for such a study can become prohibitive if many locations must be considered.

efsa JOURNAL

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: The number of monitoring sites required is discussed in Chapter 6. p50. 'Sufficient samples' How many? Ambiguity in this recommendation for assessment: 'If the concentrations are very high' How high? 'Where the exceedances just exceed the regulatory limit' Exceed by how much? 'In case of extremely highly concentrations' Exactly how high? 'Limit value' This refers to regulatory limit, do detections above LOD not demonstrate the result is due to leaching? 'Close to the limit value' How close? 'Hardly any' How many? *PPR Panel: Agreed that exact numbers are needed when the final guidance is developed. See Chapter 6.*
74	UBA (Germany)	6.1.1. General considerations	The first paragraph states that the FOCUS (2009, 2014) concept of 'one safe use' in modelling can be directly translated to monitoring data. However, chapter 9 of FOCUS (2009, 2014) remains rather vague in that aspect. The fundamental rationale at the lower tiers was to establish realistic worst-case scenarios that are protective for the diversity of agricultural areas in Europe. Each of the nine FOCUS scenarios defined by specific set of pedo-climatic parameters represents one major agricultural area in Europe. A means to convey this FOCUS climate-soil-zone approach to a suitable site selection system for targeted monitoring studies in Europe and for zonal and national risk assessments on MS level is not provided. It is, however, required to develop detailed and scientifically based guidance on how a site selection in relation to the FOCUS zone (or other new scientifically based pedo-climatic classification systems Europe) can be performed. Some recommendations are provided in the vulnerability chapter 4, but these recommendations are not related to the FOCUS zone approach. If the 'one safe use' approach was be considered in assessing GW monitoring data, our interpretation would be that at least 20–50 vulnerable monitoring sites are required for each of the FOCUS climate zones. This is also in line with the outcome of discussions on GW monitoring studies held during expert's meetings between MS and EFSA during the EU Peer Review. This number of sites was estimated to be necessary to be able to reliably conclude on the GW risk within a major agricultural region as represented by a FOCUS scenario. However, we want to point out that these numbers were not scientifically derived. We believe that answering the essential regulatory question how to translate the lower tiers FOCUS concept of 'one safe use' to the assessment of monitoring data requires profound scientific background. Further discussions and guidance on this issue are urgently needed. PPR Panel: The number of sites needed for the one-safe use approach in FOCUS groundwater is discussed in Chapter 6.
75	CropLife Europe	6.1.2. Assessment of monitoring data as a function of the exposure assessment option	A single defined protection goal across Europe is likely impossible. MS unlikely to agree with a harmonised approach. The 'Equivalent FOCUS pass' approach as a way of gaining agreement on monitoring at the EU level without having to identify a specific harmonious protection goal across EU which is likely unachievable due to different MS objectives. Individual MS can address study options appropriate for their own considerations. *PPR Panel: This issue is outside the scope of a scientific working group.*
76	CropLife Europe	6.2.3. Sites	When instrumenting new sites, groundwater flow direction needs to be understood via triangulation rather than relying on topography or potentiometric maps. Other long term established well networks in MS may not need to ascertain groundwater flow in the same was as a newly installed site, due to the legacy of data and already

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			known understanding of groundwater flow. Reference/datum points should be recorded for several permanent features of a well e.g. casing, piezometer and ground level. At commercial fields, well location may not be ideal to capture all groundwater flow and locations can be compromised to satisfy landowner's needs. *PPR Panel: See comment 51*.
77	ANSES (France)	6.2.3. Sites	comment 1: Page S51: The following is reported: 'For the selected sites, the following information should be included, when available and relevant. Please note that the kind of information required and/or available will depend on the type of monitoring study'. It would be useful to clarify which information should be provided for each type of monitoring study. comment 2: Page S51: In addition to 'Drains, possible influence of nearby surface water bodies', the impact of water sampling related to irrigation should be also reported since this can affect quickly the piezometric level and modify the normal flux of water. *PPR Panel: Agree that guidance is needed for different types of study design. Irrigation should be considered when this is normal agricultural practice as it can enhance leaching*.
78	UBA (Germany)	6.2.5. Presentation of the results	In several chapters of the paper (e.g. 3.1.1 and 4.4) the use of the 90th percentile to describe the GW risk is mentioned as an option. In our view, the appropriateness of the use of the 90th percentile is an important regulatory issue that remains open. Apart from the question how such a percentile might be calculated in practice (e.g. from all values at each site or from all sites, etc.), the protection goal and the respective study design must be defined beforehand. The harmonised agreement in the EU that the leachate concentration in the 1 m unsaturated soil zone estimated by conservative modelling over a multi-year period is used for GW risk assessment follows the precautionary principle and is considered as adequate for the protection of aquifers in large areas. Consequently, annual average concentrations in combination with the defined spatiotemporal percentiles in the lower tiers 1 and 2 can directly be compared to higher tier field experiments (FOCUS Tier-3) with a comparable exposure assessment goal (e.g. lysimeter). In contrast, it is not consistent to use the pertinent trigger values in combination with the same spatiotemporal percentiles readily for the assessment of GW monitoring results. In most cases, the filter screens of the monitoring wells will be installed at depths > 1 m in the saturated soil zone and thus in a different environmental compartment. This cannot be directly compared to the unsaturated soil zone, because e.g. the lateral flow processes which may occur in that zone need to be considered additionally. Therefore, a higher conservatism in percentage or percentile selection compared to the FOCUS specific protection goal in the lower tiers (temporal and spatial percentiles implemented by the choice of the leachate from 1 m soil depth as the decisive result to safeguard all groundwater aquifers) should be applied when GW is directly sampled. *PPR Panel: The choice of the spatiotemporal percentile in the exposure assessment goal is a risk manager's decision. However, as it is generally accepted that the concentration decreases with depth due to dispersion and/or dissipation, it is scientifically defensible to say that the concentration at 1-m depth is a conservative estimate for the groundwater concentration. Furthermore, higher percentile than 90% could challenge the tiered approach. See Chapter 3*.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
79	UBA (Germany)	6.2.6. Discussion and conclusions	We fully support the new approach of presenting the results of GW monitoring studies as the number of detections/samples (< LOD; > LOD but < parametric values; > parametric values). *PPR Panel: Agreed*. In addition to the important points already mentioned in the chapter (e.g. pressure of use, connectivity, representativeness), the discussion should contain an all-encompassing consideration of the available information on the groundwater risk posed by the respective substance (PECgw modelling, experimental leaching studies, monitoring). If the data do not provide a consistent overall picture, sound reasoning should be presented on the causes for the differing results across the tiers and/or study types. Particularly, if the monitoring results indicate lower GW concentrations, it must be reasonably justified why the results from the lower tiers (PECgw modelling, experimental leaching studies) are systematically erroneous. *PPR Panel: The Panel considers it acceptable that Tier-4 overrules the lower tiers (according to the basics of a tiered approach). Of course, this only applies if the monitoring study is well performed. Guidance is to be developed. See further Chapter 3 (principles of the tiered approach) and Chapter 6 on study design*.
80	ANSES (France)	6.2.6. Discussion and conclusions	Page S52: It is agreed that results should be presented first in the 3 classes proposed (concentrations < LOD, > LOD but < 0.1 μg/L, > 0.1 μg/L). Then, a deeper analysis should be provided, depending on the defined protection goal. At least some general recommendations could be given regarding the use of the relevant statistical parameter (maximum/mean/percentile value), depending on the number of sites, number of samples, sampling frequency, etc. *PPR Panel: Agreed. In the discussion section of any submitted study there should be a statistical paragraph*.
81	UBA (Germany)	7. Public monitoring data collected by third party organisations	According to chapter 7 in the publication Gimsing et al. (2019) the European Groundwater Monitoring Network for Regulators provides a detailed list of items how publicly available groundwater monitoring data from environmental agencies or similar public institutions in European member states should be collected and characterised. We recommend the list of items as basis for characterisation of public monitoring data that are provided in relation to active substance approvals in the EU and/or zonal and national authorisations in context of Regulation (EC) 1107/2009. *PPR Panel: The list of recommendations is added as an appendix to the Statement and could serve as a starting point for guidance development*.
82	CropLife Europe	7. Public monitoring data collected by third party organisations	The section notes a lot about what should be done and practically it should be clearer on how it should be done to promote acceptance. Shared metabolites (may arise from legacy PPPs, other PPPs and other chemicals) – there is no mention of this issue, and that some degree of pragmatism may thus be required when interpreting the data;

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			Validity of generally polluted sites within the assessment should be clarified, e.g. sites that are very polluted with an array of chemicals, that such sites are really outliers and make no contribution to an assessment of the current state of the environment; This guidance generally excludes hard rock and karstic sites from the new monitoring programme collection methodology ... what does this mean for classing and interpreting data from such sites in public monitoring programmes? There is no detail on procedures for data assessment e.g. trend assessment is mentioned but no method is outlined or referred to, like WFD guidance; age of groundwater is mentioned but no preferred techniques are described, etc; o What statistical summaries should be produced and how these summaries should be produced e.g. identifying outliers; dealing with left-censored data, different censoring levels, spatial auto-correlation, etc. While site characterisation is covered as being helpful to data interpretation there is little said about ancillary quality data being used e.g. like a site being in a designated water quality zone e.g. a nitrate vulnerable zone; or of a defined vulnerability class according to a national scheme/dataset; the use of NO_3-N or other proxies indicating pressure from agriculture and connectivity with upstream fields; example 3 of appendix 2 suggests a metabolite of a sugar beet a.i. was used for this purpose. *PPR Panel: A reflection on these issues is given in Chapter 7 of the Statement. We agree that generally more guidance is needed, including data quality requirements.*
83	UBA (Germany)	7. Public monitoring data collected by third party organisations	The uncertainties arising from the lack of data available to characterise the monitoring sites (regarding pressure of use, connectivity, groundwater flow direction, travelling times etc.) often limit the regulatory use of public monitoring data according to Regulation (EC) 1107/2009. The available information will seldomly be sufficient to fulfil all the criteria to qualify the results for a refinement with the aim to directly overwrite the results of PECGW modelling or experimental leaching studies. On the other hand, these data may contribute valuable monitoring results representing a status quo for many aquifers under a wide variety of environmental and agricultural conditions. It might be possible that results of public monitoring unveil 'blind spots' in the groundwater risk assessment that are the consequence of atypical substance properties or transport processes not covered by the leaching models. Therefore, public monitoring data should always be used alongside the typical studies submitted to meet the data requirements according to Commission Regulation (EU) 283/2013 as additional information. *PPR Panel: Agreed that public monitoring data can generally not be used for regulatory decisions according to Regulation (EC) 1107/2009. However, they should be used as additional evidence, but only if they fulfil certain data requirements according to guidance that still needs to be developed. See further Chapter 7.* In this chapter and in others the paper underlines the importance of pesticide use data. Depending on the study design and protection goal use data for the past decade(s) and in the entire catchment might be essential, but not always easily available. This is particularly the case if public monitoring data should be used. A comprehensive data base on pesticide use would be very helpful to support groundwater monitoring studies. A

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			basis for this is currently discussed within the renewal of the Statistics on Agricultural Inputs and Outputs (SAIO) legislation. Please note that useful synergies between the process of guidance development for the use of (public) groundwater monitoring data in future and the SAIO process might come up in the following discussions. *PPR Panel: Noted*.
84	CropLife Europe	7.1. Different sources, objectives, and representativeness of publicly available monitoring data	(Follow up to our comment under section 7) • There is no mention made of agencies curating their databases which may result in changes over time, e.g. when data is removed changing the max recorded concentration or a site that was once identified as an issue now being ok; numbers of samples for last year or 2 being under-represented as not all data fully loaded yet, etc. • Water companies, regions, nations do not always release the coordinates of GW abstraction sites for security reasons and if they do they often degrade the accuracy. This impacts (i) data availability and (ii) accuracy of subsequent spatial analyses; We need regulators to account for this i.e. we are not able to supply what they want as we cannot access the data; and/or they should help facilitate access so we can; • Regions/nations may not monitor for a substance as they do not consider it to be an issue; as such no data being available may not be a sign of uncertainty but rather a sign of certainty ... this needs to be factored into the assessment; • The reason for the monitoring data collection should be taken into account as part of the quality assessment e.g. incident investigation data should not be considered in general assessments; • Need it noted that the use of public data from one MS might be used and accepted in another MS if the relevance can be demonstrated. *PPR Panel: Agree that these issues should be covered in a guidance document (see Chapter 7).*
85	CropLife Europe	7.2. Other factors influencing the quality of data from official monitoring programmes	(7.1 and 7.2) When instrumenting new sites, groundwater flow direction needs to be understood via triangulation rather than relying on topography or potentiometric maps. Other long term established well networks in MS may not need to ascertain groundwater flow in the same way as a newly installed site, due to the legacy of data and already known understanding of groundwater flow. Reference/datum points should be recorded for several permanent features of a well e.g. casing, piezometer and ground level. At commercial fields, well location may not be ideal to capture all groundwater flow and locations can be compromised to satisfy landowner's needs. *PPR Panel: See comment 51*.
86	AGES (Austria)	8. Appendix 1: Protection goals	As already indicated above, a 90th spatial/temporal percentile of annual average concentrations from a groundwater monitoring study as proposed in the case of the SPG option 2 to 5 may not be acceptable at the level of Member States, where all groundwater (in space and time) may need protection. In view of AGES, it is not defensible to apply the assessment endpoint at FOCUS Tier-1 (i.e. a realistic 'worst-case' 90th percentile concentration) at FOCUS Tier-4 as well, as FOCUS Tier-1 and FOCUS Tier-4 account for water in entirely different environmental compartments with a different spatial and temporal dimension. *PPR Panel: The selection of the spatiotemporal percentile is a risk management decision because it also includes socio-economic considerations. However, selecting a percentile above the 90th-spatiotemporal percentile may*

efsa JOURNAL

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			challenge the tiered approach. This issue is discussed in Chapter 2 (protection goal) and Chapter 3 (tiered approach).
87	ANSES (France)	8. Appendix 1: Protection goals	Comment 1: It is our understanding that the spatial and/or temporal percentile(s) proposed for the definition of the protection goals are only given as example and are to be defined further by risk managers as part of overall protection goal definition. This should be more emphasised. As a general manner, the status of these options (illustrative purpose) should be clear in any further statement, to avoid that these are taken as true risk assessment options in future monitoring studies proposed by the applicants. In addition, the rationale behind the choices made for each protection goal could be further explained. *PPR Panel: The Panel agrees that a harmonised exposure assessment goal is needed before regulatory guidance can be developed. We also agree that the goals described in Gimsing et al. (*2019*), including the spatial and temporal percentiles and their combinations are examples. The Panel proposes as an alternative to address six elements of the ExAG. This is addressed in Chapter 2 of the Statement.* Comment 2: Consideration of the spatial units: it is indicated that 'the spatial units define also the area or elements over which concentrations can be averaged'. The definition of the spatial unit, and how it may influence the definition of the protection goal, is not perfectly clear. It should be further described for each option and how it is linked with the spatial statistical population of units should be clarified. *PPR Panel: See Chapter 2 for proposals on the spatial unit.* Comment 3: The sampling frequency should be considered in the question about the temporal statistical population of concentration. For example, the use of annual average concentrations may not be suitable if only 2 samplings per year are done. Please see additional comments in the attached document. *PPR Panel: The issue of sampling frequency is discussed in Chapter 6.*
88	CropLife Europe	9. Appendix 2: Examples of study designs for groundwater monitoring studies	Section 3 states that examples IV, VII and VIII are field leaching studies but in appendix 2, table 3 number V and VII are described as field leaching. This should be aligned. In my view examples IV, V, VII and VIII are field leaching studies therefore clearly illustrating the need to cover this study type in this document. *PPR Panel: The distinction between field leaching studies and monitoring studies needs to be clarified in future guidance. See Chapter 6.*
89	UBA (Germany)	9. Appendix 2: Examples of study designs for groundwater monitoring studies	The approach and rationale behind using annual median mass fluxes instead of the soil pore concentration for a comparison of relative leaching risk between regions appears plausible (in example I). Nevertheless, we do not fully agree with the reasoning that the concentration is not particularly useful in the context of comparing one area with another to estimate leaching potentials. We think that for the identification of the most vulnerable sites in terms of leaching the influence of the heterogeneous distribution of pore water volume due to varying soil types and weather conditions is the most realistic reproduction of environmental conditions. Therefore, the possible differences in the outcome of both methods should be compared and scientifically evaluated.

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			PPR Panel: A discussion on the use of annual median mass fluxes instead of the pore water concentration is discussed in Chapter 4. The Panel does not see the need for addressing both options.
90	ANSES (France)	9. Appendix 2: Examples of study designs for groundwater monitoring studies	In example VI, considering that the wells are placed at the edge of treated fields with screens at the top of the saturated zone, one could consider that exposure assessment option 2 is more suitable than exposure assessment option 4. The choice of option 4 could be further discussed. As a general manner, the exposure assessment options proposed in relation with the examples should be further explained to help the reader understand the difference between the different options. *PPR Panel: Exposure assessment options should be discussed based on the six dimensions of the ExAG as described in Chapter 2.*
91	CropLife Europe	12. Appendix 5: GIS data available at European level for vulnerability mapping	Additional spatial data: • Soil data: ISRIC Soil Grids https://soilgrids.org/ • Soil data: HWSD (Harmonised World Soil Database) (Harmonised world soil database) (Harmonised world soil database v1.2|FAO SOILS PORTAL|Food and Agriculture Organisation of the United Nations) • Soil data: OpenLandMap – OpenGeoHub Foundation: Connect|Create|Share|Repeat • Climate data: European Climate Assessment & Dataset project Home European Climate Assessment & Dataset (ecad.eu) • Climate data: World Clim Historical climate data – World Clim 1 documentation • Climate data: Land Data Assimilation System Land Data Assimilation System|LDAS (nasa.gov) • Groundwater data: Global patterns of groundwater table depth (WTD) Catalogue http://thredds-gfnl.usc.es/thredds/catalog/GLOBALWTDFTP/catalog.html • Groundwater data: Hydrogeological Map of Europe Produktcenter (bgr.de) • Groundwater data: European map of groundwater pH and calcium Zenodo (all links are in our pdf submission document) *PPR Panel: Noted.*
92	CropLife Europe	13. Appendix 6: Time of flight modelling methodology	Time of flight modelling provides a rough estimate in the absence of any other information. The time-of-flight analysis outlined in appendix 6 is an approximate methodology to get a general idea of the likely transit times under conditions of chromatographic flow, and to understand which sampling strategies might be appropriate. It is a useful methodology in the absence of any other data, as is often the case when the decision to conduct a monitoring study is made and an idea of how long it might be needed to be conducted for residues to be seen and to ensure that exceedances are not missed. Once a monitoring site has been selected, it may be that site hydrology and percolation estimates would give a more accurate picture of potential travel times. *PPR Panel: Your justification is noted.*
93	UBA (Germany)	13. Appendix 6: Time of flight modelling methodology	The prediction of the width of the solute peak as basis to determine the sampling schedule seems to be a promising way to ensure capturing the peak residues. But with the knowledge acquired while compiling active substance Assessments Reports a special focus and awareness needs to be directed at the MTC in terms of minimum time at the defined target concentration. The presented approach always should show the overall results (min., avg. and max. months at MTC) for all modelling examples shown in figure 36. Additionally, further discussion should be directed at the proclaimed 70% of the PECmax value as sufficient representation of the

Number	Organisation or name or respondent	Chapter	Comments and PPR Panel responses
			residues in groundwater. The extension of the soil profile for the modelling exercise is comprehensible but the justification considering the dispersion length enlargement for the last two soil horizons needs to be more substantial. *PPR Panel: The issues are addressed in Chapter 6.2 under the topic 'Travel time (time-of-flight modelling)'.*
94	CropLife Europe	14. Appendix 7: Examples of coupling leaching models with hydrogeological models	The example provided in a poster contribution by Sur et al. (2011) on the coupling of a leaching model with a hydrogeological model has meanwhile been published in a scientific journal article and complemented with two additional case studies. Therefore, we would like to draw the attention of reader to the following article to get a more comprehensive understanding of the matter: Herrmann, M., Sur, R. Natural attenuation along subsurface flow paths based on modelling and monitoring of a pesticide metabolite from three case studies. Environ Sci Eur 33, 59 (2021). https://doi.org/10.1186/s12302-021-00490-2 *PPR Panel: Noted*.