Mechanistic Effect Modeling of Earthworms in the Context of Pesticide Risk Assessment: Synthesis of the FORESEE Workshop

Valery E Forbes,* Annika Agatz,† Roman Ashauer,§ Kevin R Butt,¶ Yvan Capowiez,** Sabine Duquesne,†† Gregor Ernst,‡‡ Andreas Focks,¶¶ Andre Gergs,§§ Mark E Hodson,|| Martin Holmstrøp,*** Alice SA Johnston,**** Mattia Meli,***** Dirk Nickisch,****** Silvia Pieper,****** Kim J Rakel,******* Melissa Reed,******** Joerg Roembke,********* Ralf B Schäfer,********** Pernille Thorbek,*********** David J Spurgeon,************ Erik Van den Berg,************* Cornelis AM Van Gestel,************** Mathilde I Zorn,*************** and Vanessa Roeben,***********************

*Department of Ecology, Evolution and Behavior, University of Minnesota, St Paul, Minnesota, USA
†ibacon GmbH, Roßdorf, Germany
‡Syngenta Crop Protection AG, Basel, Switzerland
§Department of Environment and Geography, University of York, York, United Kingdom
∥INRAE, UMR 1114 EMMAH, INRA/Université d’Avignon, Domaine Saint Paul, Agroparc, Avignon, France
¶¶UBA Umweltbundesamt, FGIV-1.3, Section Plant Protection Products, Dessau, Germany
 §§Bayer AG, CropScience Division, Monheim, Germany
 §§§Wageningen Environmental Research, Wageningen, the Netherlands
 #Department of Bioscience, Aarhus University, Silkeborg, Denmark
 ||School of Biological Sciences, University of Reading, Reading, United Kingdom
 ††Adama Agricultural Solutions Ltd, Suresnes Cedex, France
 ‡‡‡RIFCON GmbH, Hirschberg, Germany
 ‡‡§gaiac, Aachen, Germany
 ‡‡¶Health and Safety Executive, York, United Kingdom
 ‡‡¶¶ECT Oekotoxikologie GmbH, Floersheim, Germany
 ‡¶¶Institute for Environmental Sciences, University of Koblenz and Landau, Landau, Germany
 §§§§BASF SE, APIDEE, Limburgerhof, Germany
 §§§§§UK Centre for Ecology and Hydrology, MacLean Building, Wallingford, Oxon, United Kingdom
 §§§§§Department of Ecological Science, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
 §§§§§§CTGB, Ede, the Netherlands

ABSTRACT

Earthworms are important ecosystem engineers, and assessment of the risk of plant protection products toward them is part of the European environmental risk assessment (ERA). In the current ERA scheme, exposure and effects are represented simplistically and are not well integrated, resulting in uncertainty when the results are applied to ecosystems. Modeling offers a powerful tool to integrate the effects observed in lower tier laboratory studies with the environmental conditions under which exposure is expected in the field. This paper provides a summary of the (In)Field Organism Risk modEling by coupling Soil Exposure and Effect (FORESEE) Workshop held 28–30 January 2020 in Düsseldorf, Germany. This workshop focused on toxicokinetic–toxicodynamic (TKTD) and population modeling of earthworms in the context of ERA. The goal was to bring together scientists from different stakeholder groups to discuss the current state of soil invertebrate modeling and to explore how earthworm modeling could be applied to risk assessments, in particular how the different model outputs can be used in the tiered ERA approach. In support of these goals, the workshop aimed at addressing the requirements and concerns of the different stakeholder groups to support further model development. The modeling approach included 4 submodules to cover the most relevant processes for earthworm risk assessment: environment, behavior (feeding, vertical movement), TKTD, and population. Four workgroups examined different aspects of the model with relevance for risk assessment, earthworm ecology, uptake routes, and cross-species extrapolation and model testing. Here, we present the perspectives of each workgroup and highlight how the collaborative effort of participants from multidisciplinary backgrounds helped to establish common ground. In addition, we provide a list of recommendations for how earthworm TKTD modeling could address some of the uncertainties in current risk assessments for plant protection products. Integr Environ Assess Manag 2021;17:352–363. © 2020 SETAC

Keywords: Cross-species extrapolation Plant protection products Population modeling Soil organisms Uptake routes

* Address correspondence to veforbes@umn.edu
Published 10 September 2020 on wileyonlinelibrary.com/journal/ieam.
BACKGROUND

Earthworms are important ecosystem engineers that increase soil fertility and provide a wide range of ecosystem services (Blouin et al. 2013). They are included in the safety assessment of pesticides in the European Union (EU), which is prescribed by European legislation (Regulation EC 1107/2009 [EU 2011] and guidance document SANCO/10329/2002) (SANCO 2002). In the EU, pesticides can be authorized only if no unacceptable effects on nontarget organisms, biodiversity, or the ecosystem will occur. In the risk assessment procedures, testing representative species of earthworms and assessing the risks to this group is deemed to cover soil-inhabiting Oligochaeta, belonging to the families Lumbricidae (earthworms) and Enchytraeidae (potworms). Species from both groups are used in standard ecotoxicological tests, but only tests with lumbricid earthworms need to be submitted according to EU data requirements for pesticides (283/2013 and 284/213) (EC 2013a, 2013b). The annex to the data requirement mentions as relevant Organisation for Economic Co-operation and Development (OECD) guidelines for testing the genus Eisenia (e.g., E. fetida, E. andrei) (OECD 1984, 2016).

In 2017, the European Food Safety Authority (EFSA) requested an opinion from the Panel on Plant Protection Products and their Residues (PPR) on the science behind the risk assessment of plant protection products for in-soil organisms in preparation for a guidance update (EFSA PPR 2017). In this opinion, a review was presented on the current risk assessment scheme, and proposals were made for further progress. The use of mechanistic effect models was suggested in this opinion, for example, models clarifying the relationships between internal concentrations and toxicological effects over time for endogeic earthworms. However, gaps in the currently available models were identified, as well as a need for research on their applicability domains in soil risk assessment. Therefore, a workshop was organized to clarify these issues. The present synthesis provides a summary of the workshop findings and recommendations.

The current soil risk assessment follows a tiered approach, starting with simple assumptions with effects characterized in standardized laboratory studies. In addition, whereas the exposure assessment can take into account spatiotemporal variability of pesticides in soil, abiotic parameters (such as soil temperature and moisture), and soil composition, the effect assessment for earthworms is based on the outcome of a reproduction test that does not take such factors into account. In the first tier, a chronic earthworm study (guide line 222, OECD 2016) is required, which has the aim of assessing the intrinsic toxicity of the tested substance. Exposure is assumed to vary in time because the pesticides interact with the soil and degrade, and hence a constant exposure cannot necessarily be maintained. In this type of study, adult E. fetida are exposed to a series of pesticide concentrations, and the relevant endpoints are assessed only once during the study period, that is, after 28 d for survival and growth and after 56 d for the number of juveniles. Therefore, only a limited mechanistic understanding of the underlying effect is provided. The results of this test, expressed in terms of a no observed effect concentration (NOEC) for mortality, reproduction, and growth or a 10% effect concentration (LC10/EC10) for mortality and reproduction is then related to a worst case predicted environmental concentration (PEC) for soil to obtain a Tier 1 estimate of risk. Should the ratio of toxicity to exposure (TER) be below a defined trigger value (currently 5 for chronic risk assessment in the EU; Regulation EU 546/2011 [EU 2011]), a risk is indicated, and a higher tier assessment (intended to include more realism in exposure and/or effects; Solomon et al. 2008), such as a field study with earthworms (ISO 2014) can be performed to refine the risk.

The limited conceptual integration of exposure and effect assessments in Tier 1 soil risk assessments leads to uncertainty when the results are extrapolated to different ecosystems. Furthermore, there is a large gap between a simple Tier 1 laboratory study and a full field study, which suggests the need for intermediate tiers between these options. Key uncertainties that could be reduced through the use of intermediate tier assessments include addressing actual exposure profiles arising from earthworm vertical movement and spatiotemporal variability in pesticide concentration, and extrapolation of the results from field studies to other environmental and/or agricultural situations beyond actual conditions in the field study. A critical source of uncertainty is the possible difference in sensitivity between species tested in the laboratory and species found in the field. The low field relevance of Eisenia species has been recognized, and suggestions to use Aporrectodea caliginosa as an additional test species are underway (e.g., Bart et al. 2018). Decreasing the uncertainties in the risk assessment requires a better understanding of the risks of pesticides to different species of earthworms, including whether there is any relationship between quantifiable traits and toxicological sensitivity.

Mechanistic effect modeling offers a potentially powerful tool to integrate pesticide exposure and effects and extrapolate results observed in lower tier laboratory studies to exposure scenarios that are expected in the field (EFSA PPR 2017). In particular, individual-based models (IBMs) are of interest because they allow for a high degree of realism, can help to quantify uncertainties, and can integrate processes that occur across multiple scales (DeAngelis and Grimm 2014). Populations are represented as consisting of discrete individuals, and population-level behavior and effects emerge from interactions of the individuals with each other and with their environment (DeAngelis and Grimm 2014). Johnston, Hodson et al. (2014), Johnston, Holmstrup et al. (2014), and Johnston et al. (2015, 2018) developed several earthworm IBMs that incorporate realistic earthworm behavior, address spatiotemporal variability in pesticide exposure, and integrate exposure and effects using an energy budget approach. Given the growing recognition of the power of mechanistic effect models for use in ERA (Hommen et al. 2016) and recent regulatory guidance on their development, testing, and documentation (EFSA 2014; EFSA PPR 2017), there are clear opportunities to address specific
issues identified in ERA approaches by developing dedicated models aiming at appropriate tiers of the ERA process. With that said, it is clearly impractical to develop and apply unique effect models and related behavior and exposure scenarios for every single species. Instead, an approach in which selected earthworm species can represent broader ecological groups of earthworms is needed. In addition, acceptance of models and exposure scenarios for ERA will be facilitated through consistent and transparent procedures for the development and use of effect models, species behavior, and exposure scenarios. The acceptance can be further promoted by evaluation and documentation and through broad stakeholder buy-in (Forbes et al. 2019).

**OBJECTIVE AND RATIONALE**

The (In)Field Organism Risk modEling by coupling Soil Exposure and Effect (FORESEE) Workshop was held 28–30 January 2020, in Düsseldorf, Germany. The overall focus of the workshop was to bring together scientists from different stakeholder groups (i.e., regulatory authorities, industry, contract research organizations [CROs]), and academic scientists to discuss the current state of earthworm modeling. The workshop aimed to identify research gaps and explore how mechanistic effect modeling of earthworms could be applied to soil organism risk assessments. In particular, we considered how the different model outputs could be used in the regulatory framework and in the tiered approach prescribed by the recent EFSA scientific opinion that addresses the state of the science on risk assessment of plant protection products for in-soil organisms (EFSA PPR 2017). In support of these goals, the workshop aimed to discuss a common modeling framework for earthworms and address the requirements and concerns of the involved stakeholder groups at an early stage of model development.

Johnston, Hodson et al. (2014), Johnston, Holmstrup et al. (2014), and Johnston et al. (2015, 2018) developed and validated a suite of earthworm models that integrate exposure, effects, energy budgets, behavior (movement), and life cycles. Based on these models, Roeben et al. (2020) initiated the development of a modular framework for earthworm modeling (FORESEE) that aims to cover most of the relevant earthworm ecological categories (i.e., ecotypes). The FORESEE model is mechanistic and aims to provide spatio-temporal realism in earthworm behavior, as well as exposure and effects of pesticides. The workshop was based on the FORESEE modeling approach, containing 4 submodules to cover relevant aspects of earthworm modeling: environment, behavior (feeding and movement), toxicokinetics–toxicodynamics (TKTD), and population dynamics. In practical terms, the environment module is linked to an IBM containing movement, TKTD, and population submodels from which earthworm population dynamics emerge. The environment module utilizes outputs from pesticide exposure models (e.g., PEARL [Leistra et al. 2001], PELMO [Klein 1991], HYDRUS [Simunek et al. 2012]), providing spatially and temporally explicit information on soil moisture, temperature, organic matter content, bulk density, and total and porewater pesticide concentrations. The behavior module simulates the feeding and vertical movement of different species representing 4 major ecological categories of earthworms using a trait-based approach. The TKTD module covers the toxicity of pesticides to earthworms using the General Unified Threshold model of Survival (GUTS) (Jager et al. 2011; Jager and Aschauer 2018) for lethal effects and Dynamic Energy Budget (DEB)-TKTD (Jager et al. 2006; Jager 2019) models for sublethal effects. The population module incorporates existing population models of different species (e.g., Johnston, Hodson et al. 2014; Johnston, Holmstrup et al. 2014; Johnston et al. 2015, 2018). Before use in regulatory risk assessments, all modules should be evaluated independently and be designed to allow for updating when additional knowledge becomes available (EFSA 2014). Preferably, the evaluation, release, and version control of the effect model versions could take place within the already existing framework for the version control of pesticide fate models, that is, the EFSA Chaired FOCUS Version Control Group.

During the workshop, participants from academia, regulatory authorities, CROs, and industry were divided into 4 workgroups. These workgroups examined different parts of FORESEE and addressed various questions relevant to earthworm mechanistic effect modeling: risk assessment, earthworm ecology, uptake routes, extrapolation and testing against experimental data sets, and ecotoxicological study needs and data gaps.

Workgroup (WG) 1 focused on how model outputs could fit into future risk assessment procedures for earthworms. Participants discussed how the ecotoxicity assessment and fate inputs fit into the modeling approach. Furthermore, they explored how FORESEE outputs could be used to refine the risk assessment of earthworms in different ways and how the modeling approach fits into the tiered ERA employed under EU regulatory requirements.

Workgroup 2 focused on earthworm ecology. The group discussed the main factors governing the behavior of important ecological groups of earthworms in arable soils and whether their movement could be described by a set of behavioral traits. Furthermore, the workgroup looked at other traits, such as reproduction, vertical distribution, and feeding type, and how such traits likely influence the movement of earthworms. The modelers confronted model assumptions with knowledge on earthworm ecology provided by the rest of the group, and in this way, tested whether the model is sufficiently realistic while being generally applicable to different earthworm species.

Workgroup 3 focused on exposure of earthworms, uptake routes, and TKTD modeling. Participants discussed relevant pesticide exposure routes (dermal vs oral) and concentrations (porewater vs total soil or litter concentrations) of pesticides for earthworms. Their discussions included the influence on exposure of different soil properties (e.g., organic matter), model calibration with laboratory toxicity tests, and multiple pesticide applications.

Workgroup 4 focused on species extrapolation and testing to increase the validation status of the model.
The species typically used in laboratory experiments to evaluate pesticide risks often differ from those characteristic of relevant field habitats. Therefore, species extrapolation and model testing to increase the validation status were considered within the same workgroup. The workgroup discussed models to extrapolate ecotoxicological sensitivity across species and how to address data gaps. In addition, issues related to data availability and requirements for model evaluation were discussed.

In the present workshop synthesis, we present the perspectives of each workgroup and highlight how the collaborative effort involving multiple stakeholders and representing a diversity of scientific expertise was able to reach consensus on a suite of recommendations and priorities for future work to develop FORESEE into an implementable tool for pesticide risk assessment in the EU.

**KEY FINDINGS**

**Risk assessment (WG 1)**

The current risk assessment scheme (SANCO/10329/2002) has a gap between the risk assessment tiers (SANCO 2002). There are currently only a few intermediate refinements of risk (e.g., laboratory tests using natural soils or additional test species) between the Tier 1 risk assessment using the chronic laboratory study and the higher tier assessment based on a field study. We identified several levels of the risk assessment in which modeling tools can be used. In the lower tier risk assessment, a model could be used to understand the impact of soil properties and bioavailability on the toxicity to soil organisms. Likewise, a model combining the realistic movement of earthworms (e.g., in relation to soil moisture or food availability) with a spatiotemporal exposure profile could help to generate refined exposure endpoints. Those endpoints could be used to calculate a refined TER, based on the simulated movement and resulting exposure. At the next tier, a potential advance would be to combine this spatiotemporal exposure pattern with TKTD modeling following the principles outlined in EFSA’s scientific opinion on TKTD modeling (EFSA PPR 2018) to predict risk at the level of individuals. Another use of the models could be to compare the Tier 1 assessment with the field study to check on the degree of conservatism of the Tier 1 assessment. However, the current standard chronic earthworm laboratory study (OECD 2016) is mainly suitable for setting NOECs and/or EC10s but is not adequate for parameterizing effect models. The standard study does not provide information on the time course of effects and cannot differentiate between reproductive effects and mortality of newly hatched juveniles. An option could be a modified test to allow counting of cocoons and assessment of hatching rate, in addition to the direct measurement of juvenile production. Moreover, the results of the Tier 1 laboratory chronic test are based on nominal pesticide concentrations, given that there is no requirement to measure soil or tissue concentrations. Thus, for modeling to be used to refine risk assessments, new study designs that increase the number of recorded parameters are needed.

Furthermore, environmental conditions and farming practices vary across regions and crops. Modeling could facilitate the extrapolation of the findings from the conditions under which the field studies were conducted to other conditions. As for Tier 1 data, measuring the exposure profile in the field is necessary. Following successful validation, the model could then be used to extrapolate to the relevant untested conditions such as other regions, crops, good agricultural practices (GAPs), or across multiple years. Modeling could also be used to inform the revision of the risk assessment scheme. For instance, it could be used in conjunction with field studies to calibrate the lower tiers of the risk assessment or assess the relevant soil depth at which to apply the PEC to be used in standard risk assessments. Finally, modeling could be used for interpretation of field study results and exploration of mitigation and compensation options.

If exposure models are used to provide input for the modeling, the resolution of data has to be considered. Some current EU regulatory pesticide fate models include sufficient temporal and spatial resolution (e.g., FOCUS, PEARL) (Leistra et al. 2001; FOCUS 2014) but are suitable only for simulating uniform application scenarios, such as spray applications to the crop or soil surface, or applications by injection and incorporation into the soil (Van den Berg et al. 2016). If the application is not homogenous (e.g., drip application, tree row application, or precision farming), the fate models will need a higher spatial resolution to produce outputs useful for higher tier ERAs. Although high-resolution 2-dimensional fate models exist, such as HYDRUS (Šimůnek et al. 2012) or 2DROPS (Agatz and Brown 2017), they are not yet open access.

As is the case for all models used for ERA, it is critical that FORESEE is evaluated and documented thoroughly following the principles of good modeling practice as recommended by EFSA (2014). Evaluation options include testing against additional laboratory and field studies, sensitivity and robustness analysis (i.e., pushing the limits of the model and testing its domain of applicability), evaluation of sub-models in fit-for-purpose studies, and using results of control and toxic standard treatments from field studies. In addition, an uncertainty analysis of assessments based on the model, model assumptions, and parameterizations would need to be included, also in comparison to standard assessment procedures (EFSA PPR 2018; EFSA 2019).

Scenarios need to be clearly defined to represent relevant environmental conditions, and the fate models must provide the necessary inputs for temperature and soil moisture. However, scenarios that have been chosen to be worst case from a pesticide fate perspective may not be worst case from an ecological perspective (e.g., if dry conditions during the exposure window keep the earthworms in deeper soil layers). Thus, the fate scenarios need to be evaluated to ascertain whether they are sufficiently worst case from an ecological perspective to determine whether new scenarios are needed. Models are acknowledged as a useful tool for understanding processes or simulating effects that cannot...
Earthworms are often categorized into 3 ecotypes: epigeic, endogeic, and anecic (Bouché 1977; Bottinelli et al. 2020). Epigeic (surface-living) and anecic (vertical burrowing) earthworms rely on leaf litter at the soil surface for habitat (epigeic only) and food, whereas geophagous endogeic earthworms live in temporary horizontal burrows in the mineral soil (Jégou et al. 1998; Capowiez et al. 2014). Distinct patterns of movement and surface activity across earthworm ecological groups, together with the environmental fate of different pesticide applications, can strongly influence pesticide exposure through the soil profile. Accurate assessments of pesticide effects on earthworm populations necessitate the consideration of each ecological group (Tomlin 1992). Ecotypes might not always explain the behavior observed in the field, but are currently the most accepted concept and therefore chosen as model categories. Differences in reaction to changes in environmental parameters might also be observed not only between species but also between juvenile and adult worms of 1 species, leading to different movement ranges and distribution patterns over time.

Eisenia fetida, Aporrectodea caliginosa, and Lumbricus terrestris are often mentioned as representative species of epigeics, endogeics, and anecics, respectively (Lee 1985). However, the position of L. terrestris within these ecological categories, as defined by Bouché (1977), has been questioned. Many authors refer to L. terrestris as epi-anecic rather than anecic (e.g., Hoeffner et al. 2019) due to differences in diet and behavior. Epi-anecic species first build a burrow or shelter and subsequently use it to forage at the soil surface, whereas anecic species (also referred to as “true anecic”) burrow more continuously in the soil and thus ingest more soil (Ferrière 1980; Bastardie et al. 2005). To fully represent earthworm ecotypes (and thus communities) currently found in agricultural lands, the workgroup decided that for a model to be applied in soil risk assessment, 4 main ecotypes are needed: epigeic, endogeic, epi-anecic, and anecic. Different species can be used to represent these ecotypes, and good examples are Lumbricus castaneus for epigeics, A. caliginosa for endogeics, L. terrestris for epi-anecics, and Aporrectodea nocturna or Aporrectodea longa for anecics. Eisenia fetida is also used to represent epigeics, primarily because so much data are available for this species.
Besides the abiotic factors listed above, other factors, including exposure to pesticides, food availability, avoidance behavior, interspecies interactions, spatial competition, and intraspecies interactions, can also affect the movement of earthworms and, therefore, the risk of exposure (Capowiez and Belzunec 2001; Uvarov 2017). The workgroup discussed the possible extent of the different influences. Most of the workshop participants concluded that these features do not necessarily need to be included in the model, depending on the level of realism required to address the specific question at hand and considering trade-offs between generality and realism. However, some participants recommended that the relevance of these features ought to be analyzed in a sensitivity analysis prior to considering trade-offs between simplification and realism. If identified as important, they should be considered for inclusion in the model to increase reliability of model outputs.

In some cases, there may be a lack of available data, which has to be acknowledged when choosing a modeling approach. For the epigeic and endogeic ecotypes, it is assumed that mating takes place when earthworms meet another individual randomly within the soil. For epi-anecic species, foraging is the primary driver for movement on the soil surface, and if 2 adult individuals meet, the earthworms may mate and reproduce, depending on the season.

The workgroup concluded that the development of a trait-based approach for the movement of earthworms is possible but data intensive. A list of necessary data and traits may be required for a trait-based movement model to be implemented have been identified. For some categories, the data gaps are greater than others (Table 1), but population models are available for 3 of the 4 groups, whereas a model for “true” anecics still needs to be developed (Johnston, Hodson et al. 2014; Johnston, Holmstrup et al. 2014; Johnston et al. 2015, 2018).

**Uptake routes (WG 3)**

Extrapolation of effects from standardized laboratory toxicity tests to effects in the environment is challenging because it requires several extrapolation steps. Using a TKTD framework, in combination with soil fate modeling, allows the use of mechanistic modeling to facilitate the required extrapolations. At first, the exposure to a substance with specific physicochemical characteristics in the artificial soil used in the laboratory toxicity tests has to be translated to different real soil types in the environment. In comparison to current methods, this translation can be made more accurate by explicitly modeling the fraction of active ingredient in porewater and sorbed to particles in both systems (i.e., in the laboratory toxicity test and in the environment) (Li et al. 2020).

### Table 1. Data availability, identified by Workgroup 2, for the development of realistic population modules for different earthworm species representative of 4 earthworm ecotypes

<table>
<thead>
<tr>
<th>Data type</th>
<th>Lumbricus terrestris Epi-anecic</th>
<th>Aporrectodea longa Aneics</th>
<th>Aporrectodea caliginosa Endogeics</th>
<th>Eisenia fetida/Lumbricus castaneus Epigeics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy budget</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>?</td>
</tr>
<tr>
<td>Mortality rate</td>
<td>? (lit.)</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Temperature preference</td>
<td>✓ (lit.)</td>
<td>?</td>
<td>✓ (lit.)</td>
<td>✓ (lit.)</td>
</tr>
<tr>
<td>Soil water potential preference</td>
<td>✓ (lit.)</td>
<td>✓ (lit.)</td>
<td>✓ (lit.)</td>
<td>✓ (lit.)</td>
</tr>
<tr>
<td>Soil organic matter preference</td>
<td>—</td>
<td>?</td>
<td>✓</td>
<td>?</td>
</tr>
<tr>
<td>Bulk density preference</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
<td>?</td>
</tr>
<tr>
<td>Mating as surface</td>
<td>✓</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>% time at surface</td>
<td>✓</td>
<td>? (lit.)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>% time burrowing</td>
<td>✓ (lit.)</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>% time displacing</td>
<td>✓ (lit.)</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>% time inactive</td>
<td>✓ (lit.)</td>
<td>? (lit.)</td>
<td>? (lit.)</td>
<td>?</td>
</tr>
</tbody>
</table>

✓ = available data; (lit.) = data are available from the scientific literature; — = not available; ? = requires a literature review to identify whether the data are available.
Relevant chemical properties include the partitioning coefficient \( K_D \) (or the organic-C normalized variant \( K_{OC} \)), which describes the partitioning of a chemical between the water and soil phase and, therefore, its availability for transport, uptake, and subsequent effects. This partitioning is influenced by soil composition, for example, the amount of organic C, but also by the actual soil water content. For ionizable chemicals, also the pH and speciation information, such as \( pK_a \) values, are informative. Ultimately, biodegradation rate constants need to be considered because they capture the decline of chemicals.

The second extrapolation requires accounting for different uptake routes (e.g., via skin, via gut), which are of different relative importance for different earthworm species (e.g., different movement patterns, different food sources including litter). Currently, this is not explicitly accounted for in the risk assessment, though some would argue that differentiating between uptake routes in the standard ERA may not be needed if it is sufficiently conservative. This limitation could be overcome by using the internal pesticide concentrations in the earthworm. The extrapolation can be made more accurate by a 2-step TKTD approach (Ashauer and Escher 2010). In the first step, the different uptake routes are simulated to calculate the time-variable internal exposure (approximated as whole body residues). This is proposed to be done when analyzing the laboratory toxicity study, and when simulating effects in the environment. In a second step, the effects (TDs) are simulated using the internal pesticide concentration as the forcing variable. This can be done when analyzing the laboratory toxicity study to calibrate the TKTD model and when predicting effects in the environment.

For the endpoint survival, this approach is termed “full-GUTS” (Jager et al. 2011; Ashauer et al. 2016), and the same principle can be applied to DEB-TKTD to account for sublethal effects. The GUTS approach is considered ready to be used in risk assessment in the EFSA scientific opinion on TKTD for aquatic organisms, and although the DEB-TKTD modeling approach is currently limited to research applications, its potential for future use in ERA for pesticides is recognized (EFSA PPR 2018). Key aspects of TKTD modeling can be transferred from the aquatic to the terrestrial risk assessment, in particular the calculation of exposure multiplication factors (Ashauer et al. 2013; EFSA PPR 2018) as well as many recommendations for model calibration. This same document (EFSA PPR 2018) recommends strict requirements for the validation of models.

For uptake of pesticides into soil organisms, it is essential to consider bioavailability given that there are multiple compartments of the soil in which the pesticide can be present (in porewater and sorbed to soil organic matter and soil mineral particles) and bioavailable to different extents. Pesticide properties, such as partitioning coefficients or biodegradation rate constants, and soil properties, including water content, pH values, and organic C content, can influence partitioning of the pesticide in soil. These properties can result in different concentrations in porewater, sorbed to soil organic and soil mineral particles, and in the soil pore airspace and, therefore, in differences in bioavailability. Thus, pesticide exposure depends on local conditions, pesticide properties, and earthworm ecology (e.g., movement, food sources). The pesticide distribution in the soil can be modeled using fate models, but these need to be extended to include the additional effects of soil properties that influence bioavailability if internal concentrations are to be predicted. Modeling uptake from porewater and soil particles via skin and uptake from particles via the gut accounts for bioavailability (relative contributions of the different compartments) in pesticide uptake and effects.

The model by Jager et al. (2003) is a good starting point for accounting for general uptake via dermal exposure versus feeding. We are not aware of any alternative; however, additional experimental work will be required to underpin the tentative relationship between log \( K_{OW} \) and uptake rate constants established by Jager et al. (2003) with a larger database and to account for confounding factors related to bioavailability. This is important because the rate constants acquired from experiments as described by Jager et al. (2003) may depend on environmental variables, for example, soil properties, soil water content, or temperature. Thus, new experimental and data analysis protocols are needed to disentangle the influence of environmental variables and substance properties on rate constants. The approach of acquiring uptake and elimination rate constants for both exposure routes (dermal and oral) via the partition coefficient, log \( K_{OW} \), is based on only 3 example compounds with a rather high log \( K_{OW} \) and is subject to the limitations described above (dependency on experimental variables). Until this relationship is made more robust with more data covering a wider range of log \( K_{OW} \) values, and disentangled from experimental variables, it is better to measure the actual uptake rates via gut and skin for each compound under investigation and in each soil type of interest. The limitations described here can be overcome by modeling the fate and distribution of test substances in the soil of the laboratory experiment in combination with TKTD modeling. Such data analysis may be able to disentangle the influence of environmental variables, and bioavailability, on TK rate constants from their relationship with substance properties.

Specifically, there is a need for toxicity tests for more and specifically low sorbing compounds to evaluate the usefulness of the whole approach, that is, the combination of soil fate modeling with 2-step TKTD modeling. Validation experiments can include laboratory toxicity tests with different soils and compounds (to evaluate whether bioavailability is properly accounted for) as well as experiments with different exposure patterns and field studies (see also next section).

Understanding the bioavailability and actual exposure in the toxicity test used for model calibration is essential because it enables better extrapolation to different soils in the environment. Including measurements and/or model simulations of pesticide fate in the chronic earthworm study.
(OECD 2016) would be a step toward providing a more relevant exposure estimate within current testing schemes.

**Cross-species extrapolation and model testing (WG 4)**

For earthworms, as for most taxa, a major issue hampering between-species extrapolation is that the relevant field species are not tested in the laboratory on a routine basis, and species may differ in their sensitivity and traits. As a consequence, the evaluation of TKTD models and population models based on laboratory data, by comparing them with field studies, is associated with additional challenges. The suitability of models developed for a laboratory test species, such as *E. fetida*, for field species remains uncertain and may be inaccurate if species vary in inherent sensitivity and traits.

From previous studies, it has been shown that earthworms can have different inherent sensitivities to chemicals, including pesticides (Ma and Bodt 1993; De Lima e Silva et al. 2017; Römbke et al. 2017). In a metaanalysis of species sensitivity, Pelosi et al. (2014) found that reported LC50 values for more widespread and ecologically relevant earthworm taxa were, on average, significantly lower than for *E. fetida*. This finding is indicative of a systematic lower sensitivity of this widely tested species that needs to be considered in any modeling framework. Whereas this difference in sensitivity has been observed for lethal effects, little is known regarding which traits explain the differences in inherent sensitivity, and sublethal endpoints for non-standard test species are difficult to determine.

Explicitly addressing differences in species sensitivity in mechanistic effect models ideally involves the identification of potential underlying causes for cross-species differences. Several characteristics of a species related to 1) phylogeny; 2) physiology, morphology, ecology; and 3) gene and protein expression are likely to provide mechanistic explanations for sensitivity differences among species. Species sensitivity can be represented by summary statistics, like LC50 or NOEC, but approaches for predicting TKTD model parameters are more likely to succeed because the model parameters are biologically meaningful (Ashauer and Jager 2018; Gergs et al. 2019). Van den Berg et al. (2020) hypothesize that models related to physiology, morphology, and ecology exhibit the highest prediction power for TK parameters, whereas gene and protein expression models may exhibit the highest prediction power for TD parameters.

As for TK parameters, a wide range of physical and ecological traits can potentially affect exposure and uptake. These include skin and gut wall morphology and structure; the role of the gut microbiome, which varies between species; body size in relation to passive diffusion (also applicable for life stage sensitivity); gut residence times; lipid content; and metabolic capacity (Phase I, II, and III enzyme activities) of species. Some of these trait data are simple to measure and can be collected fairly easily for widespread earthworm species, whereas others will be difficult to fully characterize. In the latter cases, it may be more efficient to categorize traits relating to TK by assessment of rates based on screening metabolism of different model compounds, rather than through detailed mechanistic prediction that attempts to cover all substances.

Toxicodynamic traits that determine sensitivity include the presence, structure, and functional motif of potential molecular targets for the chemical, the extent of damage resulting from a given level of exposure, as well as repair mechanisms. Gene and protein expression-based approaches are available for the assessment of these characteristics and can identify the presence of putative target orthologues, such as with the ECOdrug tool (Verbruggen et al. 2018). For a more detailed target-specific sequence analysis, the sequence alignment to predict across-species susceptibility (SeqAPASS) tool supports orthologue identification, as well as motif and specific residue level analyses (LaLone et al. 2016). The underlying assumption inherent in these tools is that the presence of an orthologue in a species is likely to be associated with higher sensitivity. In addition, species that possess orthologues containing conserved ligand binding motifs and key residues associated with strong ligand interactions will be more sensitive than species that lack strong ligand binding domains or residues. These assumptions have been tested in a number of selected case studies (Gunnarsson et al. 2008; LaLone et al. 2017). However, the complexities of genome evolution, including gene family expansion and reduction as well as gene and even whole genome duplications, mean that these tools are still far from being at a stage in which they fully capture all TD processes that may influence sensitivity.

If sufficiently reliable models for the extrapolation of species sensitivity toward different chemicals are available, or in cases for which laboratory toxicity data are available for the parameterization of effect models for relevant field species, confidence in population models can be increased based on field toxicity trials. An example of population-level testing is reported in Johnston et al. (2018). The use of effect data for testing TKTD models derived from field studies is limited, but field data could be used to partially (e.g., initial decline in abundance) validate predictions from TKTD models such as GUTS. It is recognized that both the range and ranking of species sensitivities may vary considerably among compounds, and pragmatic approaches for dealing with this are needed.

Any move to apply mechanistic models for modeling pesticide impacts in earthworms will require a change in current testing procedures. The current chronic earthworm test involves the assessment of survival on day 28 only and measurement of reproduction at test termination (day 56). Development of process-based approaches such as TKTD models, however, requires data at a higher temporal resolution. Designs that include measurement of survival at regular times for parameterization of TKTD models, such as GUTS, are potentially easy to conduct by extending exposure time and increasing observations of mortality to at least 4 time points. A challenge for such studies with earthworms is simply that soil, unlike water, is not
transparent. Consequently, each measurement involves disturbing the test system (e.g., by hand sorting), raising the issue of stress and potential mechanical damage. Alternatively, a destructive sampling design could be used, though this would require additional replicates. For DEB model application, the slow rate of earthworm development and the extended time scale of reproduction mean that life cycle tests measuring juvenile and adult traits over time are unlikely to be feasible, and if conducted, would need to take compound fate into consideration. It should also be noted that for many field-relevant species, following individuals from birth through adulthood and reproduction is not practical. Approaches that separately measure juvenile growth and adult reproduction have been proposed and could form the basis of a suitable method for time series data collection (Van Gestel et al. 1991; Spurgeon et al. 2003). Given the intricacies of TKTD model development and parameterization for a long-lived soil-dwelling species, a further challenge is how to validate model predictions for those species. One approach is to use mechanistic measurements, such as internal concentrations or measurements of tissue “damage” (although difficult to define) for the testing of model components. However, the targeted nature of such measurements means that they may validate only 1 parameter, rather than the output of the model as a whole. Therefore, validation of TKTD model predictions on the level of survival or reproduction based on laboratory validation tests is recommended. Ideally, the conditions for experiments for validation purposes should be different from those in the calibration experiment and reflect the (regulatory) question to be addressed by the model. Examples include variation of exposure duration, spatial variation of exposure, variation in time scale (temporal extrapolation), or variation in environmental conditions such as soil properties affecting chemical fate and exposure, and earthworm movement and population structure.

PRIORITIES FOR FUTURE WORK

We conclude that a mechanistic modeling approach, linking appropriate environmental variables, reflecting defined scenarios, TKTD processes, and movement behavior, can provide realistic individual- and population-level predictions. This approach offers promise for improving scientific understanding and informing pesticide risk assessment for earthworms in the EU regulatory framework. With that said, we have identified several areas in which more work is needed to allow FORESEE to reach its full potential. Moreover, we provide several recommendations for moving this initiative forward. Filling the remaining data gaps identified by workshop participants would enable FORESEE to achieve its full potential as a tool for refining risk and to address uncertainties in the present risk assessment for earthworms exposed to pesticides. These are shown in Table 2.

RECOMMENDATIONS

We recommend the following:

- Further developing FORESEE as a mechanistic effect model that could be applied for pesticide risk assessment

---

### Table 2. Main data gaps for earthworms in the context of this workshop on soil organism pesticide risk assessments and how filling them would improve ERA

<table>
<thead>
<tr>
<th>Data gap</th>
<th>Needed for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition of realistic worst-case environmental scenarios for modeling (spatial and temporal scales, number of spatial dimensions, soil and climate variables) and establishing link to existing exposure models</td>
<td>Relevant data for FORESEE’s environment module</td>
</tr>
<tr>
<td>Intermediate measurements of survival, growth, and reproduction in chronic earthworm study</td>
<td>Time course data to parameterize GUTS or DEB-TKTD</td>
</tr>
<tr>
<td>Toxicity test results for different soils and chemicals with a range of log K_{OW} values</td>
<td>Proof of concept with a short-term benefit to the existing risk assessment as it could be used to replace the arbitrary correction factor of 2 when log K_{OW} &gt; 2</td>
</tr>
<tr>
<td>Measured dermal and oral uptake rate constants for a wide range of log K_{OW} values disentangled from experimental variables (e.g., soil type, water content)</td>
<td>Establishing the relationship between uptake rate constants and substance properties (e.g., log K_{OW}) while accounting for bioavailability</td>
</tr>
<tr>
<td>A few comprehensive studies with measurements of several state variables (e.g., concentrations in bulk soil, porewater and earthworms and toxicity, over time)</td>
<td>Better system understanding and evaluation if model complexity is appropriate</td>
</tr>
<tr>
<td>Ecological studies</td>
<td>Data on movement differences among earthworm ecological categories</td>
</tr>
<tr>
<td>Tests of inherent toxicity in multiple worm species</td>
<td>Data needed for cross-species extrapolation and to distinguish sensitivity differences from exposure differences</td>
</tr>
</tbody>
</table>

DEB = Dynamic Energy Budget; ERA = environmental risk assessment; FORESEE = (In)Field Organism Risk modEling by coupling Soil Exposure and Effect; GUTS = General Unified Threshold model of Survival; TKTD = toxicokinetic-toxidynamic.
and parameterized for relevant earthworm ecotypes represented in European agricultural systems.

- Additional assessment of the differences in species sensitivity between standard test species and more ecologically relevant earthworm species for different compounds, given that species sensitivity can vary between chemicals.
- Further investigation of the relevance of impacts of abiotic and biotic factors on the movement of earthworms.
- Employing a data-informed, trait-based approach to simulate a set of representative earthworm species using a framework considering 4 ecotypes, which we believe to be sufficient for capturing earthworm behavioral traits regarding movement patterns in the soil. Traits to include describe moving and burrowing behavior and niche characteristics (i.e., tolerance to drought, temperature, soil bulk density, and food conditions).
- Modeling and/or measuring internal concentrations (body residues) as a step to account for different routes of uptake (e.g., dermal, gut) as a refined option in the tiered risk assessment scheme.
- Measurements of organism size, mortality, and reproductive output at intermediate time points in laboratory toxicity studies to facilitate parameterization of TKTD modeling (GUTS and DEB-TKTD). This will require a reassessment of the standardized approach currently used for earthworm toxicity testing, especially for the measurement of reproduction for intermediate time points. In addition, substantial additional effort will be involved because either soil will need to be changed at each sampling or additional replicates will be needed to allow destructive sampling.
- Further developments of FORESEE for possible use in EU pesticide risk assessment following EFSA’s guidance for good modeling practice, including detailed and transparent model documentation. This includes the consideration of model uncertainty.
- Version control of effect models in order for them to be used in the EU registration procedure. Version control can be done within the existing EFSA-Chaired Version Control Workgroup for pesticide fate models.
- Organization of a follow-up working group or targeted workshop to establish detailed experimental designs for robust model calibration and evaluation.
- Broad stakeholder engagement to achieve agreement on the data sets that FORESEE should be tested against, validation of study designs, and other criteria for model evaluation to increase the validation status of the effect models.
- Broad scientific discussion to gain consensus on appropriate ecological scenarios in which to assess risk using FORESEE given that scenarios used to derive worst-case pesticide fate estimates may not be appropriate for modeling earthworm risk.
- European Food Safety Authority to critically consider the key findings and recommendations from this workshop together with other relevant reports or published scientific information during the revision of their guidance for risk assessment of soil organisms to improve the linkage of exposure and effects and address other knowledge gaps in current ERA practice.

- Establishing a formal and transparent mechanism to encourage that models for pesticide risk assessment in the EU can be effectively and efficiently evaluated.

We acknowledge that current approaches to pesticide risk assessment include uncertainties with regard to spatio-temporal variation in pedological, climatic, and biological conditions, agronomical practices, and complexities occurring at the landscape scale (Topping et al. 2020). However, we conclude that mechanistic effect modeling of the kind described here can help to quantify and reduce uncertainties in ERA by providing improved integration of exposure and effects and by incorporating different pesticide application scenarios and greater ecological realism.

Acknowledgment—This workshop was financially supported by the European Crop Protection Association (ECPA) and organized by gaia—Research Institute for Ecosystem Analysis and Assessment (affiliated with the RWTH Aachen University). Organizing committee members are VE Forbes, R Ashauer, Y Capowiez, G Ernst, A Gergs, M Meli, KJ Rakel, RB Schäfer, P Thorbek, and V Roeben.

Disclaimer—The peer review for this article was managed by the Editorial Board without the involvement of VE Forbes.

Data Availability Statement—There are no data associated with this submission, which is a workshop summary.

ORCID
Kevin R Butt http://orcid.org/0000-0003-0886-7795
Andreas Focks http://orcid.org/0000-0002-9031-0808

REFERENCES


