Title: Mechanistic effect modeling of earthworms in the context of pesticide risk assessment: Synthesis of the FORESEE Workshop.

Authors: *Forbes Valery E[a], Agatz Annika[b], *Ashauer Roman[c,d], Butt Kevin R[e], *Capowiez Yvan[f], Duquesne Sabine[g], *Ernst Gregor[h], Focks Andreas[i], *Gergs Andre[h], Hodson Mark E[j], Holmstrup Martin[k], Johnston Alice SA[l],*Meli Mattia[m], Nickisch Dirk[n], Pieper Silvia[g], *Rakel Kim J[o], Reed Melissa[p], Roembke Joerg[q],*Schäfer Ralf B[r], *Thorbek Pernille[s], Spurgeon, David J[t], Van den Berg Erik[i], Van Gestel Cornelis AM[u], Zorn Mathilde I[v], *Roeben Vanessa[o]

Affiliations:
[a]Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, MN, 55108, USA
[b]ibacon GmbH, Roßdorf, Germany
[c]Syngenta Crop Protection AG, Rosentalstrasse 67, Basel, CH-4002, Switzerland
[d]Department of Environment and Geography, University of York, Wentworth Way, Heslington, York, YO10 5NG, UK
[e]School of Forensic and Applied Sciences, University of Central Lancashire, Preston, PR1 2HE, UK
[f]INRAE, UMR 1114 EMMAH, INRA/université d'Avignon, Domaine Sait paul, Agroparc, 84914 Avignon, France
[g]UBA Umweltbundesamt, FGIV- 1.3, Section "Plant Protection Products", Wörlitzer Platz 1, 06844 Dessau, Germany
[h]Bayer AG, CropScience Division, Alfred-Nobel-Straße 50, 40789 Monheim, German
[i]Wageningen Environmental Research, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[j]Department of Environment and Geography, University of York, Heslington, York, YO10 5NG, UK
[k]Department of Bioscience, Aarhus University, Vejlsøvej 25, DK-8600 Silkeborg, Denmark
[l]School of Biological Sciences, University of Reading, Reading, RG6 6AH, UK
[m]Adama Agricultural Solutions Ltd, 33 rue de Verdun, 92156 Suresnes Cedex, France
[n]RIFCON GmbH, Goldbeckstraße 13, 69493 Hirschberg, Germany
[o]gaiac, Kackertstr. 10, 52074 Aachen, Germany
[p]Health and Safety Executive, Mallard House, Kings Poll, York, UK
[q]ECT Oekotoxikologie GmbH, Boettgerstr. 2-14, D-65439 Floersheim, Germany
[r]Institute for Environmental Sciences, University of Koblenz and Landau, Landau, Germany
[s]BASF SE, APD/EE, APD/EE, Speyerer Strasse 2, 67117 Limburgerhof, Germany
[t]UK Centre for Ecology and Hydrology, MacLean Buidling, Wallingford, Oxon, OX10 8BB, UK
[u]Department of Ecological Science, Faculty of Science, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands
[v]CTGB, Bennekomseweg 41, 6717 LL Ede,The Netherlands

*Indicates member of the Organizing Committee.

Data availability statement: There are no data associated with this submission which is a workshop summary.

Acknowledgments – This workshop was financially supported by the European Crop Protection Association (ECPA) and organized by gaiac – Research Institute for Ecosystem Analysis and Assessment (affiliated with the RWTH Aachen University).

Abstract: Earthworms are important ecosystem engineers, and assessment of the risk of plant protection products towards 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 applying the results 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 FORESEE Workshop ((In)Field Organism Risk modEling by coupling Soil Exposure and Effect) held January 28-30, 2020 in Düsseldorf, Germany. This workshop focussed on toxicokinetic-toxicodynamic (TKTD) and population modeling of earthworms in the context of environmental risk assessment. The goal was to bring together scientists from different stakeholder groups to discuss the current state of soil invertebrate modeling, explore how earthworm modeling could be applied to risk assessments, and 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 four 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.

Key Words: cross-species extrapolation; plant protection products; population modeling; soil organisms; uptake routes

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) No. 1107/2009 and guidance document SANCO/10329/2002). In the EU, pesticides can only be authorized if no unacceptable effects on non-target 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). The Annex to the data requirement mentions as relevant OECD guidelines for testing the genus *Eisenia* (e.g., *E. fetida*, *E. andrei*).

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. This 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 (GD 222, OECD 2017) is required, which has the aim of assessing the intrinsic toxicity of the tested substance. Exposure is assumed to vary in time, since 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, i.e., after 28 days for survival and growth, and after 56 days 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 ten percent Effect Concentration (LC_{10}/EC_{10}) 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, 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 11268-3) 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 extrapolating the results 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, as 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 et al. (2014a, 2014b, 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 environmental risk assessment (ERA) (Hommen et al., 2016) and recent regulatory guidance on their development, testing, and documentation (EFSA PPR 2014, 2017), there are clear opportunities to address specific issues identified in ERA approaches by developing dedicated models aiding 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 FORESEE ((In)Field Organism Risk modEling by coupling Soil Exposure and Effect) Workshop was held January 28 – 30, 2020, in Düsseldorf, Germany. The overall focus of the workshop was to bring together scientists from different stakeholder groups (i.e., regulatory authorities, industry, and 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 addressing 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 et al. (2014a, 2014b, 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). FORESEE is mechanistic and aims to provide spatiotemporal realism in earthworm behavior, as well as exposure and effects of pesticides. The workshop was based on the FORESEE modeling approach containing four 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, PELMO, HYDRUS), 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 four 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 Ashauer 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 et al. 2014a, 2014b, 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, i.e., the EFSA Chaired FOCUS Version Control Group.

During the workshop, participants from academia, regulatory authorities, CROs, and industry were divided into four 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 datasets, and ecotoxicological study needs and data gaps.

Workgroup 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 this 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. 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 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 2017) is mainly suitable for setting NOECs and/or EC_{10}s 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, as there is no requirement to measure soil or tissue concentrations. So 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) but are only suitable 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 two-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 submodels 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 2018; 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 be tested in the laboratory, such as effects of repeated exposure over multiple years or extrapolation to other GAPs. Furthermore, they can help to calibrate the risk assessment, e.g., from Tier 1 to reference tier (field), as well as for refining the risk estimates and addressing uncertainties associated with realistic conditions when Tier 1 ERA identifies a non-acceptable risk.

Earthworm Ecology and Behavior (WG 2)

Earthworm vertical movement plays an important role in population-level exposure to pesticides in the field. Understanding how different earthworm species move in response to environmental changes is crucial for effective risk assessments of pesticides in soils. In general, earthworm movement is determined by the ecological category to which they belong and various abiotic and biotic factors (Roeben et al. 2020).

Earthworms are often categorized into three 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 one 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, four 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. *E. fetida* is also used to represent epigeics, primarily because so much data are available for this species.

The reliability of population models fundamentally depends on the availability of data, and the suitability of different earthworm species for population models depends on the ecotypes most at risk.

Individual-based models, for instance, require detailed information on the biology and behavior of the modeled species at the individual- and population levels, both for model development and model validation. Across earthworm ecotypes, abiotic factors play an important role in driving movement behavior, and thus also the possible exposure to a pesticide. Four environmental variables have been identified as critical and feasible to be used for simulating the behavior and vertical movement of earthworms: soil water potential (Gerard 1967; Holmstrup 2001), soil organic matter content (Le Couteulx et al. 2015; Frazao et al. 2019), temperature (Eriksen-Hamel and Whalen 2006), and bulk density of the soil (Kretzschmar 1991). For anecic and epi-anecic species, light is an additional factor to be considered (Nuutinen et al. 2014). Roeben et al. (2020) provide a detailed review of the effects of biotic and abiotic factors on the vertical movement of earthworms.

IBMs have the advantage of allowing interaction between an individual and its virtual environment. The identified environmental variables that influence the movement of earthworms are stored in the modeling environment, which is represented through patches in a spatially-explicit setting. One approach to incorporate the simultaneous influence of these four environmental variables on earthworm movement in an IBM is a patch quality index. This index is also part of the modeling environment and determines the movement decisions of individuals in the movement module. The index scales the attractiveness of each soil patch according to the four variables from 1 (attractive) to 0 (not attractive). In this way, the quality index can account for the combined effects of temperature, soil water potential, organic matter content, and bulk density on earthworm movement. To be able to represent the different ecotypes and their preferences realistically, the workgroup suggested a dynamic trade-off between the following factors:

- If temperature and water potential are within a defined performance range, organic matter content should be most important.
- If temperature and water potential are outside this range, organic matter content has no importance.
- The last factor is the bulk density, which inhibits the movement of earthworms with increasing density.

The performance ranges, slopes, and threshold values depend on the ecotype/species modeled and should be fitted to laboratory and field data individually. Thereby, the patch quality index can cover the different importance of factors for different ecotypes.

Besides the abiotic factors listed above, other factors, including exposure to pesticides, food availability, avoidance behavior, inter-species interactions, spatial competition, and intra-species interactions, can also affect the movement of earthworms and, therefore the risk of exposure (Uvarov 2017; Capowiez and Belzunces 2001). 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 two 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 existing knowledge gaps for representative earthworms from the given ecological categories can be found in Table 1. Ideally, an energy-budget model is available for the representative species of each ecotype. Furthermore, data on mortality and longevity of the species are needed and how these traits are influenced by abundance. Moreover, information about preferences towards the four environmental factors determining movement is essential. Finally, information on behavioral aspects is necessary, such as the percentage of time spent on different activities. This includes the time spent foraging at the surface, burrowing, moving in existing burrows, and being inactive. It is crucial to be aware that these traits and preferences can change with developmental stage and exposure, and juveniles will likely have different traits than adults. For the four ecotypes, knowledge gaps that have to be filled 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 three of the four groups, whereas a model for "true" anecics still needs to be developed (Johnston et al. 2014a, 2014b, 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).

Relevant chemical properties include the partitioning coefficient K_D (or the organic-carbon 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 carbon, but also by the actual soil water content. For ionizable chemicals, also the pH and speciation information, such as pKa values, are informative. Ultimately, biodegradation rate constants need to be considered as 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 two-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 (toxicodynamics) 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. GUTS 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 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 2018) as well as many recommendations for model calibration. This same document (EFSA 2018) recommends strict requirements for the validation of models.

For uptake of pesticides into soil organisms, it is essential to consider bioavailability as 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 carbon 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 at al. (2003) may depend on environmental variables, e.g., 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 three 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, i.e., the combination of soil fate modeling with two-step TKTD modeling. Validation experiments can include laboratory toxicity tests with different soils and compounds (to evaluate if 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 2017) would be a step towards 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 meta-analysis of species sensitivity, Pelosi et al. (2014) found that reported LC_{50} 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 a) phylogeny, b) physiology, morphology, ecology, and c) gene and protein expression, are likely to provide mechanistic explanations for sensitivity differences among species. Species sensitivity can be represented by summary statistics, like LC_{50} or NOEC, but approaches for predicting TKTD model parameters are more likely to succeed as 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.

TD 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 SEQUAPASS 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 towards 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 four 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 timescale 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 only validate one 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:

- Further developing FORESEE as a mechanistic effect model that could be applied for pesticide risk assessment 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, as 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 four 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 as 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.
- EFSA 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 ensure 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 spatiotemporal 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.

References

Agatz A, Brown CD. 2017. Introducing the 2-DROPS model for two-dimensional simulation of crop roots and pesticide within the soil-root zone. *Sci Total Environ* 586: 966-975.

Ashauer R, Escher BI. 2010. Advantages of toxicokinetic and toxicodynamic modelling in aquatic ecotoxicology and risk assessment. *J Environ Monit* 12: 2056 - 2061.

Ashauer R, Thorbek P, Warinton JS, Wheeler JR, Maund S. 2013. A method to predict and understand fish survival under dynamic chemical stress using standard ecotoxicity data. *Environ Toxicol Chem* 23: 954–965.

Ashauer R, Albert C, Augustine S, Cedergreen N, Charles S, Ducrot V, Focks A, Gabsi F, Gergs A, Goussen B, Jager T, Kramer NI, Nyman A.-M, Poulsen V, Reichenberger S, Schäfer RB, Van den Brink PJ, Veltman K, Vogel S, Zimmer EI, Preuss TG. 2016. Modelling survival: exposure pattern, species sensitivity and uncertainty. *Scientific Reports* 6: 29178, doi:10.1038/srep29178.

Ashauer R, Jager T. 2018. Physiological modes of action across species and toxicants: the key to predictive ecotoxicology. *Environ Sci: Processes Impacts* 20: 48-57.

Bart S, Amossé J, Lowe CN, Mougin C, Pery ARR, Pelosi C. 2018. *Aporrectodea caliginosa*, a relevant earthworm species for a posteriori pesticide risk assessment: current knowledge and recommendations for culture and experimental design. *Environ Sci Pollut Res* 25: 33867-33881.

Bastardie F, Capowiez Y, Renault P, Cluzeau D. 2005. A radio-labelled study of earthworm behaviour in artificial soil cores in term of ecological types. *Biol Fertil Soils*. 41(5):320–327. doi:10.1007/s00374-005-0847-6.

Blouin M, Hodson ME, Delgado EA, Baker G, Brussaard L, Butt KR, Dai J, Dendooven L, Peres G, Tondoh JE, et al. 2013. A review of earthworm impact on soil function and ecosystem services: Earthworm impact on ecosystem services. *Eur J Soil Sci.* 64(2):161–182. doi:10.1111/ejss.12025.

Bottinelli N, Bouquet P, Hedde M, Capowiez Y. 2020. Explicit definition of earthworm ecological categories – Marcel Bouché's triangle revisited. *Geoderma* 372, 114361; https://doi.org/10.1016/j.geoderma.2020.114361

Bouché MB. 1977. Strategies lombriciennes. *Soil Org Compon Ecosyst. Ecological Bulletins* 25:122–132.

Capowiez Y, Belzunces L. 2001. Dynamic study of the burrowing behavior of *Aporrectodea nocturna* and *Allolobophora chlorotica*: interactions between earthworms and spatial avoidance of burrows. *Biol Fertil Soils* 33: 310-316.

Capowiez Y, Sammartino S, Michel E. 2014. Burrow systems of endogeic earthworms: effects of earthworm abundance and consequences for soil water infiltration. *Pedobiologia* 57: 303-309.

DeAngelis DL, Grimm V. 2014. Individual-based models in ecology after four decades. F1000Prime Rep. 6. doi:10.12703/P6-39. [accessed 2019 Nov 11]. http://www.f1000.com/prime/reports/b/6/39.

De Lima e Silva C, Brennan N, Brouwer JM, Commandeur D, Verweij RA, van Gestel CAM. 2017. Comparative toxicity of imidacloprid and thiacloprid to different species of soil invertebrates. *Ecotoxicology* 26: 555-564.

EFSA 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, 92 pp. doi:10.2903/j.efsa.2014.3589.

EFSA 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, 225 pp. doi:10.2903/j. efsa.2017.4690.

EFSA 2018. Guidance on uncertainty analysis in scientific assessments. EFSA Journal 2018;16(1):5123, 39 pp. doi: 10.2903/j.efsa.2018.5123.

EFSA 2019. Guidance on communication of uncertainty in scientific assessments. EFSA Journal 2019;17(1):5520, 73 pp. doi: 10.2903/j.efsa.2019.5520.

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, Smith RH, Stemmer M, Sundh I, Tiktak A, Topping CJ, Wolterink G, Cedergreen N, Charles S, Focks A, Reed M, Arena M, Ippolito A, Byers H and Teodorovic I, 2018. Scientific Opinion on the state of the art of Toxicokinetic/Toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. EFSA Journal 2018; 16(8):5377, 188 pp. https://doi.org/10.2903/j.efsa.2018.5377

Eriksen-Hamel NS, Whalen JK. 2006. Growth rates of *Aporrectodea caliginosa* (Oligochaetae: Lumbricidae) as influenced by soil temperature and moisture in disturbed and undisturbed soil columns. *Pedobiologia*. 50(3):207–215. doi:10.1016/j.pedobi.2005.10.008.

EU 546/2011. Commission Regulation (EU) No 546/2011 of 10 June 2011 implementing Regulation (EC) No 1107/2009 of the European Parliament and of the Council as regards uniform principles for evaluation and authorisation of plant protection products OJ L 155, 11.6.2011, p. 127–175

Ferrière G. 1980. Fonctions des lombriciens. VII. Une méthode d'analyse de la matière organique végétale ingérée. *Pedobiologia*. 20(4):263–273.

Forbes VE, Schmolke A, Accolla C, Grimm V. 2019. A plea for consistency, transparency, and reproducibility in risk assessment effect models. *Environ Toxicol Chem* 38: 9-11.

Frazao J, de Goede R, Capowiez Y, Pulleman M. 2019. Soil structure formation and organic matter distribution as affected by earthworm species interaction and crop residue placement. *Geoderma* 338: 453-463.

Gerard BM. 1967. Factors affecting earthworms in pastures. *J Anim Ecol.* 36(1):235–252. doi:10.2307/3024.

Gergs A, Rakel KJ, Liesy D, Zenker A, Classen S. 2019. Mechanistic effect modeling approach for the extrapolation of species sensitivity. *Environ Sci Technol,* 53: 9818−9825.

Gunnarsson L, Jauhiainen A, Kristiansson E, Nerman O, Larsson DGJ. 2008. Evolutionary conservation of human drug targets in organisms used for environmental risk assessments. *Environ Sci Technol* 42: 5807-5813.

Hoeffner K, Santonja M, Cluzeau D, Monard C. 2019. Epi-anecic rather than strict-anecic earthworms enhance soil enzymatic activities. *Soil Biol Biochem.* 132:93–100. doi:10.1016/j.soilbio.2019.02.001.

Holmstrup M. 2001. Sensitivity of life history parameters in the earthworm *Aporrectodea caliginosa* to small changes in soil water potential. *Soil Biol Biochem* 33: 1217-1223.

Hommen U, Forbes VE, Grimm V, Preuss TG, Thorbek P, Ducrot V. 2016. How to use mechanistic effect models in environmental risk assessment of pesticides: Case studies and recommendations from the SETAC Workshop MODELINK. *Integr Environ Assess Manage* 12: 21-31.

ISO International Organization for Standardization 2014. Soil quality - Effects of pollutants on earthworms - Part 3: Guidance on the determination of effects in field situations", 2nd edition. ISO 11268-3, Geneva Switzerland.

Jager T, Fleuren RHLJ, Hogendoorn EA, de Korte G. 2003. Elucidating the routes of exposure for organic chemicals in the earthworm, *Eisenia andrei* (Oligochaeta). *Environ Sci Technol* 37(15):3399–3404. doi:10.1021/es0340578.

Jager T, Heugens EHW, Kooijman SALM. 2006. Making sense of ecotoxicological test results: Towards application of process-based models. *Ecotoxicology* 15: 305-314.

Jager T, Albert C, Preuss TG, Ashauer R. 2011. General unified threshold model of survival - a toxicokinetic-toxicodynamic framework for ecotoxicology. *Environ Sci Technol* 45: 2529-2540.

Jager T, Ashauer R. 2018. Modelling survival under chemical stress. A comprehensive guide to the GUTS framework. Toxicodynamics Ltd. York, UK, p. 193. https://leanpub.com/guts_book

Jager T. 2019. Making sense of chemical stress. Application of Dynamic Energy Budget theory in ecotoxicology and stress ecology. Version 2.0. Leanpub https://leanpub.com/debtox_book, Amsterdam.

Jégou D, Cluzeau D, Wolf HJ, Gandon Y, Trehen P. 1998. Assessment of the burrow system of *Lumbricus terrestris*, *Aporrectodea giardi* and *Aporrectodea caliginosa* using X-ray computed tomography. *Biol Fertil Soils* 26:116–121

Johnston ASA, Hodson ME, Thorbek P, Alvarez T, Sibly RM. 2014a. An energy budget agent-based model of earthworm populations and its application to study the effects of pesticides. *Ecol Model* 280:5–17.

Johnston ASA, Holmstrup M, Hodson ME, Thorbek P, Alvarez T, Sibly RM. 2014b. Earthworm distribution and abundance predicted by a process-based model. *Appl Soil Ecol* 84:112–123.

Johnston A, Sibly RM, Hodson ME, Alvarez T, Thorbek P. 2015. Effects of agricultural management practices on earthworm populations and crop yield: validation and application of a mechanistic modelling approach. *J Appl Ecol* 52:1334–1342.

Johnston ASA, Sibly RM, Thorbek P. 2018. Forecasting tillage and soil warming effects on earthworm populations. *J Appl Ecol* 55:1498–1509.

Kretzschmar A. 1991. Burrowing ability of the earthworm *Aporrectodea longa* limited by soil compaction and water potential. *Biol Fertil Soils.* 11(1):48–51. doi:10.1007/BF00335834.

LaLone CA, Villeneuve DL, Lyons D, Helgen HW, Robinson SL, Swintek JA, Saari TW, Ankley GT, 2016. Sequence alignment to predict across species susceptibility (SeqAPASS): A web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicol Sci* 153: 228-245.

LaLone CA, Ankley GT, Belanger SE, Embry MR, Hodges G, Knapen D, Munn S, Perkins EJ, Rudd MA, Villeneuve DL, Whelan M, Willett C, Zhang XW, Hecker M, 2017. Advancing the Adverse Outcome Pathway Framework-An International Horizon Scanning Approach. *Environ Toxicol Chem* 36: 1411-1421.

Le Couteulx A, Wolf C, Hallaire V, Pérès G. 2015. Burrowing and casting activities of three endogeic earthworm species affected by organic matter location. *Pedobiologia*. 58(2–3):97–103. doi:10.1016/j.pedobi.2015.04.004.

Lee KE. 1985. Earthworms: their ecology and relationships with soils and land use. Academic Press Inc. [accessed 2019 Nov 8]. https://www.cabdirect.org/cabdirect/abstract/19851998411.

Li J, Carter LJ., Boxall ABA. 2020 Evaluation and development of models for estimating the sorption behaviour of pharmaceuticals in soils. *Journal of hazardous materials* 392: 122469

Ma W, Bodt J. 1993. Differences in toxicity of the insecticide chlorpyrifos to six species of earthworms (Oligochaeta, Lumbricidae) in standardized soil tests. *Bull Environ Contam Toxicol.* 50(6). doi:10.1007/BF00209951. [accessed 2020 May 11]. http://link.springer.com/10.1007/BF00209951.

Nuutinen V, Butt KR, Jauhiainen L, Shipitalo MJ, Sirén T. 2014. Dew-worms in white nights: High-latitude light constrains earthworm (*Lumbricus terrestris*) behaviour at the soil surface. *Soil Biol Biochem.* 72:66–74. doi:10.1016/j.soilbio.2014.01.023.

OECD Organization for Economic Co-operation and Development 2004. Test No. 222: Earthworm reproduction test (*Eisenia fetida*/*Eisenia andrei*). https://doi.org/10.1787/9789264264496-en

Pelosi C, Barot S, Capowiez Y, Hedde M, Vandenbulcke F. 2014. Pesticides and earthworms. A review. *Agron Sustain Dev.* 34(1):199–228. doi:10.1007/s13593-013-0151-z.

Regulation (EC) 1107/2009, 2009. Regulation (EC) no 1107/2009 of the European Parliament and of the Council of 21 October 2009 Concerning the Placing of Plant Protection Products on the Market and Repealing Council Directives 79/117/EEC and 91/414/EEC. 50.

Roeben V, Oberdoerster S, Rakel KJ, Liesy D, Capowiez Y, Ernst G, Preuss TG, Gergs A, Oberdoerster C. 2020. Towards a spatiotemporally explicit toxicokinetic-toxicodynamic model for earthworm toxicity. *Sci Total Environ* 722: 137673 https://doi.org/10.1016/j.scitotenv.2020.137673

Römbke J, Schmelz RM, Pelosi C, 2017. Effects of organic pesticides on enchytraeids (Oligochaeta) in agroecosystems: laboratory and higher-tier tests. *Front Environ Sci* 5: 20-23.

Šimůnek J, Van Genuchten MT, Šejna M. 2012. The HYDRUS software package for simulating the two-and three-dimensional movement of water, heat, and multiple solutes in variably-saturated porous media. Technical Manual, Version. 2, p. 258.

Solomon K, Brock T, De Zwart D, Dyer SD, Postthuma L, Richards S, Sanderson H, Sibley P, van den Brink P. 2008. Extrapolation practice for ecotoxicological effect characterization of chemicals. Eds Solomon K, Brock T, De Zwart D, Dyer SD, Postthuma L, Richards S, Sanderson H, Sibley P, Van Den Brink P. CRC Press, Boca Raton, FL, 408 pp.

Spurgeon DJ, Weeks JM, Van Gestel CAM. 2003. A summary of eleven years progress in earthworm ecotoxicology. *Pedobiologia.* 47(5–6):588–606. doi:10.1078/0031-4056-00234.

Tomlin AD. 1992. Behavior as a source of earthworm susceptibility to ecotoxicants. In: Sheppard S.C. et al. (Eds) Advances in earthworm ecotoxicology. Intercept, Andover, PP 116-125.

Topping CJ, Aldrich A, Berny P. 2020. Overhaul environmental risk assessment for pesticides. *Science* 367: 360-363.

Uvarov AV. 2017. Density-dependent responses in some common lumbricid species. *Pedobiologia* 61: 1-8.

Van den Berg F, Tiktak A, Boesten JJTI, Van der Linden AMA. 2016. PEARL model for pesticide behaviour and emissions in soil-plant systems. WOt-technical report 61. Statutory Research Tasks Unit for Nature & the Environment , Wageningen, 134 pp. https://edepot.wur.nl/377664

Van den Berg SJP, Maltby L, Sinclair T, Liang R, van den Brink PJ. 2020. Cross-species extrapolation of chemical sensitivity. *Sci Total Environ* (in press). https://doi.org/10.1016/j.scitotenv.2020.141800

Van Gestel CAM, van Dis WA, Dirven-Van Breemen EM, Sparenburg PM, Baerselman R. 1991. Influence of cadmium, copper and pentachlorophenol on growth and sexual development of *Eisenia andrei* (Oligochaeta: Annelida). *Biol Fertil Soil* 12: 117-121.

Verbruggen B, Gunnarsson L, Kristiansson E, Osterlund T, Owen SF, Snape JR, Tyler CR. 2018. ECOdrug: a database connecting drugs and conservation of their targets across species. *Nucl Acids Res* 46: D930-D936.

Table 1. Data availability, identified by work group 2, for the development of realistic population modules for different earthworm species representative of four earthworm ecotypes. Ticks represent available data, (lit.) indicates that the data are available from the scientific literature, - indicates not available, and question marks require a literature review to identify whether the data are available.

	Lumbricus terrestris	*Aporrectodea longa*	*Aporrectodea caliginosa*	*Eisenia fetida / Lumbricus castaneus*	
	epi-anecic	anecics	endogeics	epigeics	
Energy budget	√	√	√	√	-
Mortality rate	?	?	?	?	?
Temperature preference	√ (lit.)	?	√ (lit.)	√ (lit.)	?
Soil water potential preference	√ (lit.)	√ (lit.)	√ (lit.)	√ (lit.)	?
Soil organic matter preference	-	?	√	-	-
Bulk density preference	√	?	√	-	-
Mating as surface	√	-	-	√	√
% time at surface	√	?	√	√	√
% time burrowing	√	? (lit.)	√	√	√
% time displacing	√	? (lit.)	√	√	√
% time inactive	√	? (lit.)	? (lit.)	?	?

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.

Data Gap	Needed for
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	Relevant data for FORESEE's environment module
Intermediate measurements of survival, growth, and reproduction in chronic earthworm study	Time course data to parameterize GUTS or DEB-TKTD
Toxicity test results for different soils and chemicals with a range of Log K_{ow} values	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 $K_{ow} > 2$
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)	Establishing the relationship between uptake rate constants and substance properties (e.g. log K_{ow}) whilst accounting for bioavailability
A few comprehensive studies with measurements of several state variables (e.g., concentrations in bulk soil, porewater and earthworms & toxicity, over time)	Better system understanding and evaluation if model complexity is appropriate
Ecological studies	Data on movement differences among earthworm ecological categories
Tests of inherent toxicity in multiple worm species	Data needed for cross-species extrapolation and to distinguish sensitivity differences from exposure differences