SCIENTIFIC OPINION

Scientific Opinion on the development of a soil ecoregions concept using distribution data on invertebrates[1]

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

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

A modelling approach for defining soil ecoregions within Europe was developed to improve the realism of exposure scenarios for plant protection products. Biological information on four soil animal groups (earthworms, enchytraeids, collembolans and isopods) was used to assign each species to different life forms, representing depth horizons in which they occur. Based on information from three countries covering a North-South gradient (Finland, Germany, Portugal), species presence-absence data were modelled using pedological and climatological information. With a triangular diagram it was possible to visualise life-form distributions for the organisms groups. Ecoregion maps were produced for earthworms and enchytraeids and revealed marked differences between the three countries. The information on the spatial distribution of the dominance classes could be transformed into depth profiles for any ecotoxicologically relevant concentration to be modelled. This procedure allows defining realistic “worst” case exposure depth profiles for risk assessment. The approach could be extended to the entire EU territory, provided more biogeographical data are available. A better resolution might also be achieved by adjusting the – presently geometrically derived – sizes of the 7 classes in the triangular diagrams. For an improved risk assessment of plant protection products it is recommended to extend exposure modelling to the litter layer. This approach would imply a refinement of the environmental risk assessment of plant protection products which needs to be discussed with risk managers. For most of the situations in Europe, the worst case soil depth profile for short term risk assessment would be litter (if present) or 0 to 1 cm depth instead of the currently used 0 to 5 cm depth.

KEY WORDS

soil, exposure in soil, soil organisms, invertebrates, environmental risk assessment, pesticides, plant protection products

---

1 On request from EFSA, Question No EFSA-Q-2009-00859, adopted on 22 September 2010.

2 Panel members : Jos Boesten, Claudia Bolognesi, Theo Brock, Ettore Capri, Anthony Hardy, Andrew Hart, Karen Hirsch-Ernst, Susanne Hougaard Bennekou, Michael Klein, Robert Luttik, Angelo Moretto, Bernadette Ossendorp, Annette Petersen, Yolanda Pico, Andreas Schäffer, Paulo Sousa, Walter Steurbaut, Anita Stromberg, Maria Tasheva, Ton Van der Linden, Christiane Vleminckx. Correspondence: ppr@efsa.europa.eu

3 Acknowledgement: The Panel wishes to thank the members of the Working Group on Ecoregions for the preparatory work on this scientific opinion: Ettore Capri, Juliane Filser, Ciro Gardi, Robert Luttik, Jörg Römbke, Andreas Schaeffer and Paulo Sousa as well as EFSA’s staff members Olaf Mosbach-Schulz and Karin M. Nienstedt for the support provided to this scientific opinion.

Suggested citation: EFSA Panel on Plant Protection Products and their Residues (PPR); Scientific Opinion on the development of a soil ecoregions concept using distribution data on invertebrates. EFSA Journal 2010;8(10):1820. [77 pp.] doi:10.2903/j.efsa.2010.1820. Available online: www.efsa.europa.eu/efsajournal.htm

SUMMARY

The European Food Safety Authority (EFSA) asked the Panel on Plant Protection Products and their Residues (PPR) to further develop the concept of soil ecoregions in the context of the revision of the Guidance Document on Terrestrial Ecotoxicology (EFSA-Q-2009-00002).

A modelling approach for defining soil ecoregions within Europe was developed to improve the realism of exposure scenarios for plant protection products. Biological information on four soil animal groups (earthworms, enchytraeids, collembolans and isopods) was used to assign each species to different life forms, representing depth horizons in which they occur. Based on information from three countries covering a North-South gradient (Finland, Germany, Portugal), species presence-absence data were modelled using pedological and climatological information. With a triangular diagram it was possible to visualise life-form distributions for all organisms groups and to identify factors determining their distribution. Ecoregion maps were produced for earthworms and enchytraeids for most of the countries and revealed marked differences between the countries. Maps are not predictive on a local scale, but give a probability of the soil biota community to be found on a regional scale. The main results obtained are:

- Maps based on modelled information are in line with ecological and biogeographical information for the organism groups considered.
- Factors determining the distribution of the organisms could be identified, in particular for earthworms and enchytraeids.
- Differences could be observed between the three countries in community composition based on life form groups of earthworms and enchytraeids.
- The transformation of the information on the spatial distribution of the dominance classes into depth profiles for any ecotoxicologically relevant concentration (ERC) to be modelled was possible.

This procedure allows defining realistic “worst” case soil depth profiles for short term exposure. For realistic worst cases, depth profiles of “litter or 1 cm” should be used for litter dwelling or epigeic organisms. For other life forms or for long term exposure, other depth profiles may represent the realistic worst case situation. For refined risk assessments the geographical variation in depth profiles, crop and soil management information, as well as data about ecology of soil organisms (e.g. different dominance distribution of soil communities) could be considered.

Provided that comparable information from other EU countries is available, the approach could be extended to the entire EU territory. However, several limitations and gaps in the availability of data constrained the analysis undertaken here:

- In comparison to grassland and forest sites, only a small number of studies in arable areas were available.
- The incomplete pedological description of the sampling sites led to the use of surrogate information on soil and climatic characteristics from additional databases, resulting in a mismatch of the resolution scales between biological and environmental data.
- The patchiness of the sampling sites led to situations where the information for different geographical areas was not homogeneous.
- Information on the abundances of species, which is more informative than presence / absence data used in the approach here, is limited.

A more complete data set would allow using more sophisticated models. In particular, a systematic data collection, the incorporation of biological knowledge on the influence of soil and climatic

conditions on the occurrence of particular species (e.g. using non-linear functions), and the use of abundance data should significantly improve the outcome of this kind of analysis.

An additional outcome was the identification of research needs addressing specific aspects of the approach described in this scientific opinion:

- While the visualization by community triangles is recommended in general, it should be checked whether the outcome can be improved, e.g. by adjusting the size of the 7 classes.
- Modelling of exposure is so far limited to the occurrence of a plant protection product in depth layers of the mineral soil. However, certain organisms may not be directly exposed via the soil. This applies in particular to the litter layer.
- Since Collembola are known to be good indicators for soil conditions on a small scale, it should be studied if data sets collected on a smaller scale would improve the model fit, e.g. the data collected in several EU research projects such as the VULCAN project[4].
- Biological monitoring programs with standardised sampling methods focussing on agricultural areas should be performed on a regular basis, partly in order to fill data gaps in certain regions but also to get better information concerning the “normal” ranges of species numbers, species composition, abundances and biomass (for general recommendations concerning biological soil monitoring programs see Römbke and Breure, 2005). This will allow the development of improved predictive models and reduce uncertainties.
- There is a general need to compile information on the biogeography and species composition of other soil organism groups not covered so far, e.g. mites, centipedes, diplopods, molluscs, nematodes and micro-organisms.
- When sampling soil organisms, it is recommended to also measure the most important soil and site parameters like pH, soil organic carbon, texture, cation exchange capacity and water-holding capacity as well as climatic factors (for details, see recommendations given in existing field test guidelines).
- More detailed geographical information on the extent of crop type and crop management practices such as tillage and/or irrigation is needed for risk assessment. However, the way to integrate this kind of information with the biogeographical information needs to be developed.

The original aim of this mandate was to define ecoregions by combining geographical information on different taxa. It was possible to define ecoregion maps for earthworms and enchytraeids for most of the model countries. The PPR Panel has demonstrated that the development of ecoregions is possible, provided that biogeographical data are available. The implementation of this approach would imply a refinement of the environmental risk assessment of plant protection products in soil which may need to be considered in the update of the Terrestrial Ecotoxicology Guidance Document (EFSA-Q-2009-00002) and, if appropriate, be discussed with risk managers. In particular, it seems that for most of the situations in Europe, the worst case soil depth profile for short term risk assessment would be litter (if present), or 0 to 1 cm depth instead of the currently used 0 to 5 cm depth. It needs to be decided if the development of ecoregions should be expanded to cover the whole EU on the medium term, to improve the risk assessment in soil. This would imply producing and gathering further biogeographical data. Further, the inclusion of additional taxa, in particularly vulnerable species, could be considered.

[4] VULCAN - Vulnerability assessment of shrubland ecosystems in Europe under climatic changes. EU FP5 Contract EVK2-CT-2000-00094

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

In the EFSA-PPR-Opinion "*The usefulness of total concentrations and pore water concentrations of pesticides in soil as metrics for the assessment of ecotoxicological effects"* (The EFSA Journal, 2009, 922: 1-90), the Panel stated that exposure assessments in soil could be refined based on a novel underlying concept using soil ecoregion maps to define ecologically relevant exposure scenarios. These soil ecoregion maps would be possible to construct based on the composition of soil organism communities (incorporating ecological and biogeographical aspects) allying information on climate and soil properties. The Panel also defined three hypothetical regions (Finland, Germany, Portugal), which represent one of the highest North-South gradient in Europe, reflecting thus the most diverse climate conditions, and which could as a consequence be used for developing the concept. Additionally, a first outline of soil ecoregions was presented and welcomed in the stakeholder workshop IRIS (*Improved Realism in Soil Risk Assessment (IRIS) - How will pesticide risk assessment in soil be tackled tomorrow?*), organized by PPR and held in Ispra in May 2009. The corresponding EFSA-report was published in July 2009.

Additionally, in the context of EFSAs mandate received early 2009 for the revision of the Guidance Document (GD) Terrestrial Ecotoxicology SANCO/10329/2002 (EFSA-Q-2009-0002), where the PPR-Panel is currently working on, the development of adequate and worst case exposure assessments in soil are crucial and the development of the soil ecoregions concept is important for defining these exposure assessments.

Therefore, in order to give continuity to the work the PPR-Panel has done so far, and considering the scheduled work on its agenda for the next few years, the development of the soil ecoregions concept represents a milestone which needs to be defined soon.

TERMS OF REFERENCE AS PROVIDED BY EFSA

The Scientific Panel on Plant Protection Products and their Residues (PPR Panel) of EFSA is asked to further develop the concept of **soil ecoregions** in the context of the revision of the Guidance Document on Terrestrial Risk Assessment (EFSA Q 2009 00002).

ASSESSMENT

1. Introduction

As put forward in the PPR Opinion "The usefulness of total concentrations and pore water concentrations of pesticides in soil as metrics for the assessment of ecotoxicological effects" (EFSA, 2009), risk assessment of plant protection products in soil would benefit from a more realistic exposure assessment. The current terrestrial risk assessment of plant protection products is performed assuming a common exposure scenario for the entire EU based on the total concentration in the top 5 cm of soil (Directive 91/414/EEC[5] ; Directive 97/57/EC[6]; CSTEE, 2000; EPPO, 2003). However, new approaches are under discussion. The new regulation on Plant Protection Products (Regulation (EC) No 1107/2009[7]) considers the definition of three regulatory zones (North, Centre and South). Moreover the PPR Opinion "Scientific opinion on outline proposals for assessment of exposure of organisms to substances in soil" (EFSA, 2010a), advises that PECsoil should be modelled at different soil depths (1, 2.5, 5 and 20 cm, based on the vertical stratification of soil organisms) in each one of the 3 regulatory zones. However, aiming at modelling the ecotoxicologically relevant concentrations in soil[8] (ERCsoil, Boesten et al., 2007), additional information is needed. To know which PECsoil value (at what depth or depths) should be considered for risk calculations, exposure scenarios based on ecological knowledge of soil communities should be developed.

Soil biota is considered to constitute a large part of the world's biodiversity. Soil organisms govern the main soil functions such as the cycling of organic matter and nutrients, the creation and stabilisation of soil structure and porosity or the degradation of pollutants, including pesticides, being directly or indirectly involved in the provision of several ecosystem services by soil (Lavelle and Spain, 2001; de Bello et al., 2010). Soil organisms comprise plant roots, micro-organisms (fungi, bacteria, algae and lichens) and soil animals. The base of the complex soil food web are carbon compounds which are either exudated by roots or enter the soil as dead organic matter. These compounds are transformed in manifold ways by micro-organisms, to be finally released as gaseous compounds into the atmosphere or as mineral salts which are taken up by plant roots. Plant roots and micro-organisms clearly dominate the living biomass in soil whereas animals interact with them and support their functions, e.g. by increasing the surface of litter for microbial growth through comminution and fragmentation, by distribution of materials and organisms or by creating pores for plant roots to grow into. Soil animals are grouped according to their size: microfauna (protists and nematodes smaller than 0.1 mm), mesofauna (e.g., collembola, mites, enchytraeids, mostly ranging between 0.1 and 2 mm) and macrofauna (larger than 2 mm, e.g., earthworms, isopods and small vertebrates).

Soil is thus a living substrate, consisting of mineral particles and organic matter in the solid and dissolved state, water, air and organisms. Different land-use types, under varying climatic and soil conditions support specific organism communities which are mainly responsible for many ecological services provided by the soil. These specific soil organism communities, as well as their ecological

---

[5] Council Directive 91/414/EEC of 15 July 1991 concerning the placing of plant protection products on the market, OJ L230, 19.8.1991,

[5] Directive 97/57/EC establishing Annex VI to Directive 91/414/EEC concerning the placing of plant protection products on the market OJ L265, 27.9.1997, p. 87-109.

[7] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L309, p. 1-50

[8] Ecotoxicologically Relevant Concentration: Interface between effect assessment and exposure assessment defined as the type of concentration that gives the best correlation to ecotoxicological effects (Boesten et al, 2007)

functions, have to be protected for maintaining the medium and long term delivery of ecosystem services to mankind (EFSA, 2010d). However, when modelling the fate of plant protection products (PPPs) in soil, the focus is on soil and climatic properties while the habitat selection of organisms is widely ignored. Thus, the risk assessment of plant protection products would benefit from an improved exposure assessment, where modelling of exposure should therefore consider both climate and soil characteristics and the distribution in the depth profile of the associated soil organism communities.

The definition of relevant exposure scenarios should consider the environment where communities normally live or occur. This environment (named here as soil habitat) is characterised by specific properties like water holding and ion exchange capacities, pH, range of temperature as well as biotic factors which influence the availability of nutrients and organic matter. Moreover, at a particular site, specific horizons with different properties and with particular associated communities can occur. Habitat properties, including abiotic factors and all organisms, determine the bioavailability of a plant protection product in soil. Thus, the risk assessment of plant protection products should consider the habitat as setting the boundaries for the exposure assessment, i.e., soil exposure scenarios for a given region should be horizon-specific, allowing modelling the ERC in different soil layers in different regions in Europe. This should be done based on a soil ecoregion approach, taking into account the biogeographical differences within the EU in terms of soil, climate and soil organism communities.

The aim of this document is to explain the development of such an approach based on selected soil invertebrate communities (earthworms, enchytraeids (potworms), collembolans (springtails), and isopods (woodlice), see more details at Section 4.1) and the steps taken to test its validity and its practicability in routine risk assessment of plant protection products. Information on crop and soil management is not considered for this purpose. The document is structured as follows:

- Firstly, an explanation is given on why a new ecoregion typology that incorporates soil fauna distribution and ecological information is needed to define ecologically relevant exposure profiles in soil (section 2).
- Secondly, the new ecoregion concept is explained in detail, including its underlying assumptions (section 3).
- Afterwards, the steps taken to test the ecoregion concept, including the establishment of the soil fauna biogeographical database, the data treatment strategy and the results obtained with the selected model countries, are presented (section 4).
- In section 5, a few examples are presented on how this new concept could be used in day-to-day risk assessment procedures.
- Finally, open points and gaps of the concept are discussed and recommendations on future actions to be taken (namely on data gaps) for the further development of the concept are given in the conclusions and research needs.

2. Why a soil ecoregion concept?

In biogeography, area units can be classified with different levels of detail. With decreasing size the classification is as follows: Ecozone ➔ Biogeographical region ➔ Ecoregion.

An Ecozone is the largest scale biogeographic division of the earth's surface based on the historic and evolutionary distribution patterns of plants and animals. The earth's surface is currently divided into seven ecozones, including Antarctica (Figure 1). Biogeographical regions are geographical reference units for describing habitat types and species which live under similar conditions. For example, Europe can be divided into 12 biogeographical regions (Figure 2). In contrast, Ecoregions (also called Bioregions) are geographically defined areas within ecozones. Ecoregions cover relatively large areas and contain characteristic, geographically distinct assemblages of natural communities (that tend to be distinct from other ecoregions). As an example the map of the European Ecoregions is shown in Figure 3, which consists of 68 classes based on climate, soil, and vegetation data.

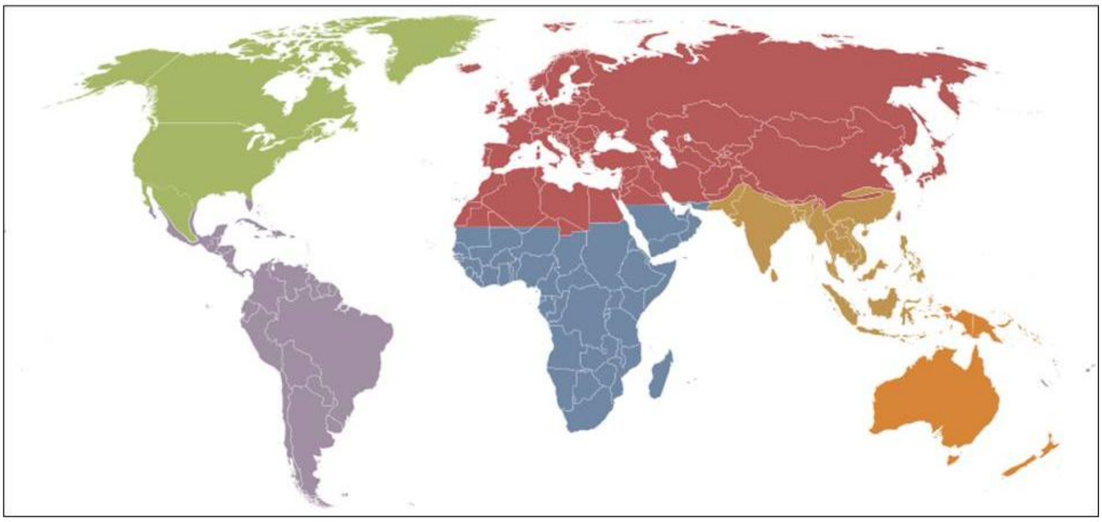


Figure 1: Ecozones of the world. Bordeaux: Palearctic; Green: Nearctic; Blue: Afrotropic; Purple: Neotropic; Brown: Indo-Malaya; Orange: Oceania (Antarctica is not represented). Source: Wikimedia (http://commons.wikimedia.org/wiki/File:Ecozones.png)

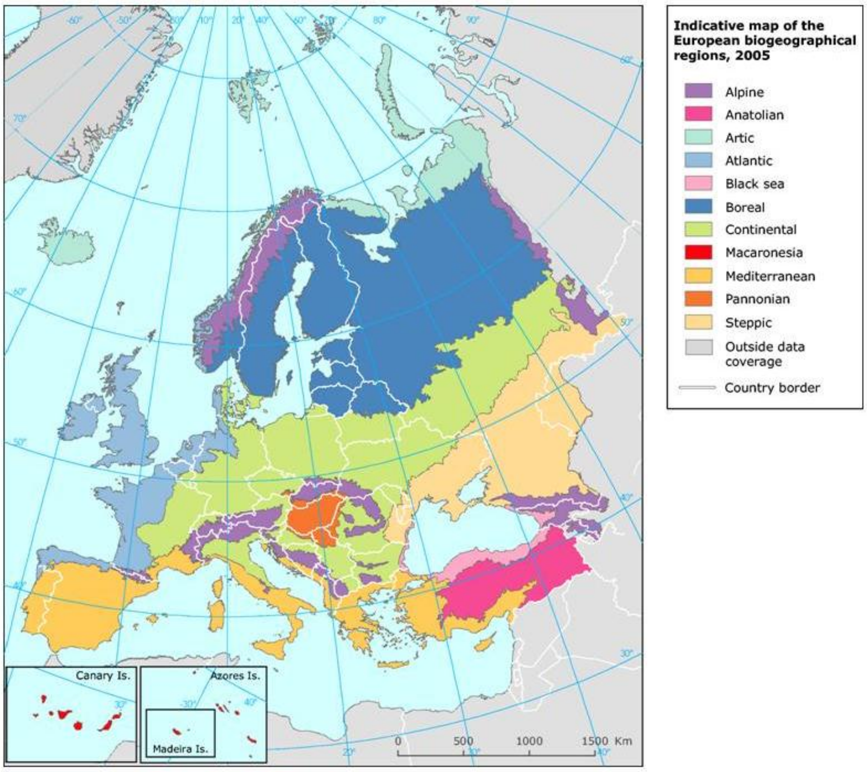


Figure 2: Biogeographical regions of Europe. Source: European Environmental Agency.

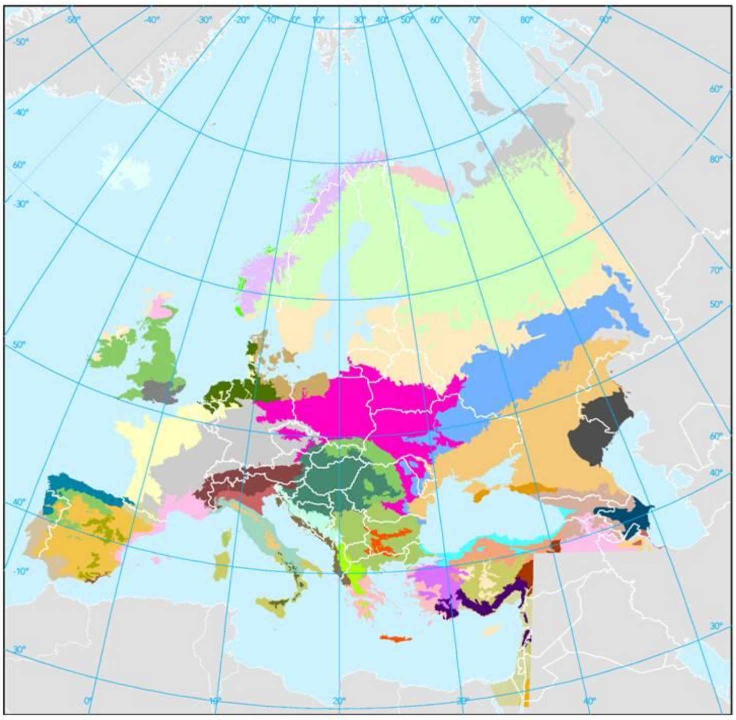


Figure 3: Ecoregions of Europe. Source: European Environmental Agency.

Ecoregions are defined by the plant cover, and usually by the Potential Natural Vegetation (PNV) which is based on different climate and, to a lesser extent, on soil or geological parameters. They do not reflect the distribution of soil invertebrates. Although distinct soil organism communities can be found in different land-use types, the presence of soil organisms is mainly influenced by soil properties like pH, organic matter and texture (Dunger, 1998; Breure et al, 2005). In fact, many species are present under very different land-uses, with different plant covers (Dunger, 1983). For example, the most widespread European earthworm species are found in different types of grasslands or forests – as long as the soil conditions, including food resources, are favourable. On the other hand, soil organisms have been called the "soil's long-term memory": due to their restricted mobility, the community composition often reflects a land-use that was given up decades ago (Dunger, 1978).

Moreover, in order to develop a typology of horizon-specific exposure profiles, it is important that the ecological information used to define the ecoregions should be based not only on taxonomic data (which species), but mainly on ecological data (which species with which characteristics), using information on the characteristics (traits) of the organisms that influence the way they are exposed to pesticides, e.g., where they live in soil. Therefore the currently available ecoregions based on plants do not allow defining exposure profiles relevant for soil organisms, e.g. relevant soil depth profiles or time windows. Thus there is a need to construct an ecoregion map based on the composition of soil organism communities. This map should incorporate taxonomic, functional and biogeographical aspects, allying information on climate, land-use and soil properties.

3. The soil ecoregion concept and underlying assumptions

The definition of soil ecoregions is partially based on a classification of soil organisms, using biological characteristics (traits) that determine the way they are exposed to plant protection products. A trait is a measurable property of organisms, usually measured at the species level and used comparatively across species[9]. Particular traits determine the exposure, e.g. over what depth the organisms are exposed, and partly the bioavailability of the chemical. An example of a trait-based classification of soil organisms is depicted in Figure 4, where different species, sharing similar traits that are indicative of the soil horizon where they live, can be grouped in "trait groups".

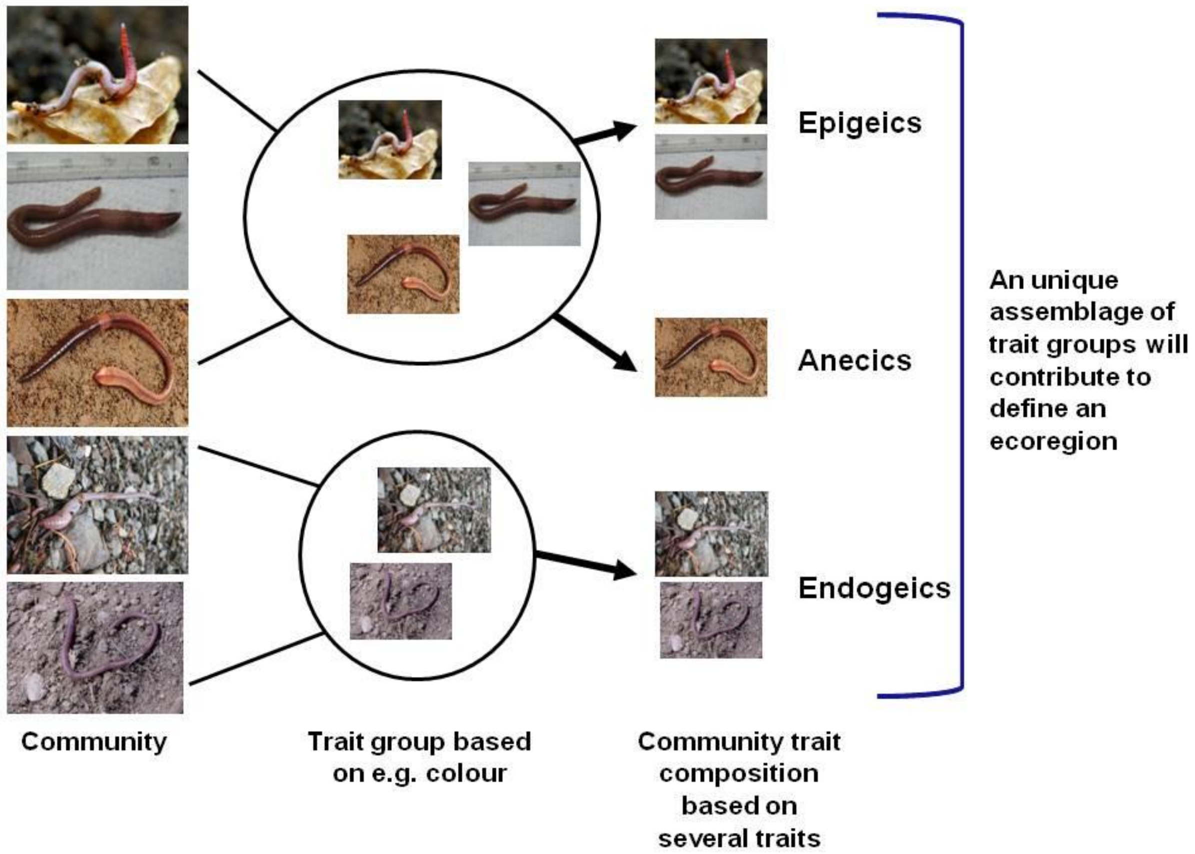


Figure 4: Example of a trait-based classification of earthworms (see Section 4.3.1). Note: **epigeic** earthworms live in the litter layer (i.e. decaying plant debris, or compost); **anecic** earthworms live in deep vertical burrows but feed at or near the soil surface and in the litter layer, especially at night; **endogeic** earthworms live and feed within the soil and burrow continuously to form a network of channels (mostly horizontal channels) in the first 20 cm of soil and around plant roots (Lavelle & Spain, 2001).

[9] Examples of traits are structural traits (e.g. permeability of exoskeleton, lipid content, and complexity of the nervous system), morphological traits (e.g. size, volume / surface ratio), physiological traits (e.g. mode of respiration, detoxifying enzymes or digestive strategy), and ecological traits (e.g. mobility, feeding behaviour, trophic level, and place in the food web).

Under this premise, the exposure assessment of plant protection products in soil could be refined based on a novel underlying concept using ecoregion maps to define ecologically relevant exposure profiles. This concept is based on the following principles:

- Europe can be divided into a number of regions defined by soil properties, land-use and climate.
- Each region supports specific soil organism communities that may play different roles in supporting relevant soil services.
- The different species within each community could be subdivided into groups based on similar traits ("trait groups") that are related in the way they are exposed to chemicals.
- The combination of soil properties, land-use, climate and the potential soil community (based on a unique assemblage of "trait groups") defines an ecoregion.
- Each ecoregion is characterised by a different set of exposure scenarios, e.g. depth profiles that are defined by the trait groups present for which homogeneous ERC values can be modelled.

A scheme representing the development of this ecoregion approach using invertebrate communities is presented in Figure 5.

In this concept, traits influencing exposure to plant protection products are the key drivers to define ecoregions and their exposure profiles. Within the soil community, it is the species' traits that determine the way they are exposed to the plant protection products. For example, species with differently structured body surfaces living in the same soil layer are exposed in different ways. However, the actual exposure/availability may differ with respect to environmental conditions, since the degradation and/or metabolisation of plant protection products as well as their availability also depend on soil properties (e.g. organic matter content) and climate (e.g. temperature). Thus, depending on the region concerned, a combination of abiotic properties (soil, climate and land-use practice) and soil communities should be considered when modelling the actual exposure to a plant protection product.

As mentioned, a specific set of exposure profiles can be described within each ecoregion, e.g. different soil horizons for which homogeneous ERC values can be modelled. When modelling the ERC of plant protection product at a specific site, the result is not only relevant for that specific set of profiles and that specific site but for all sites with comparable combinations of specific abiotic and biotic factors, i.e., sites belonging to the same ecoregion. In particular, the different life forms of the organism groups assessed in this opinion, are exposed in different soil depth profiles, as shown in Table 1. According to our concept, different ecoregions can be characterised by a similar set of exposure profiles, but the ERC values modelled will be different.

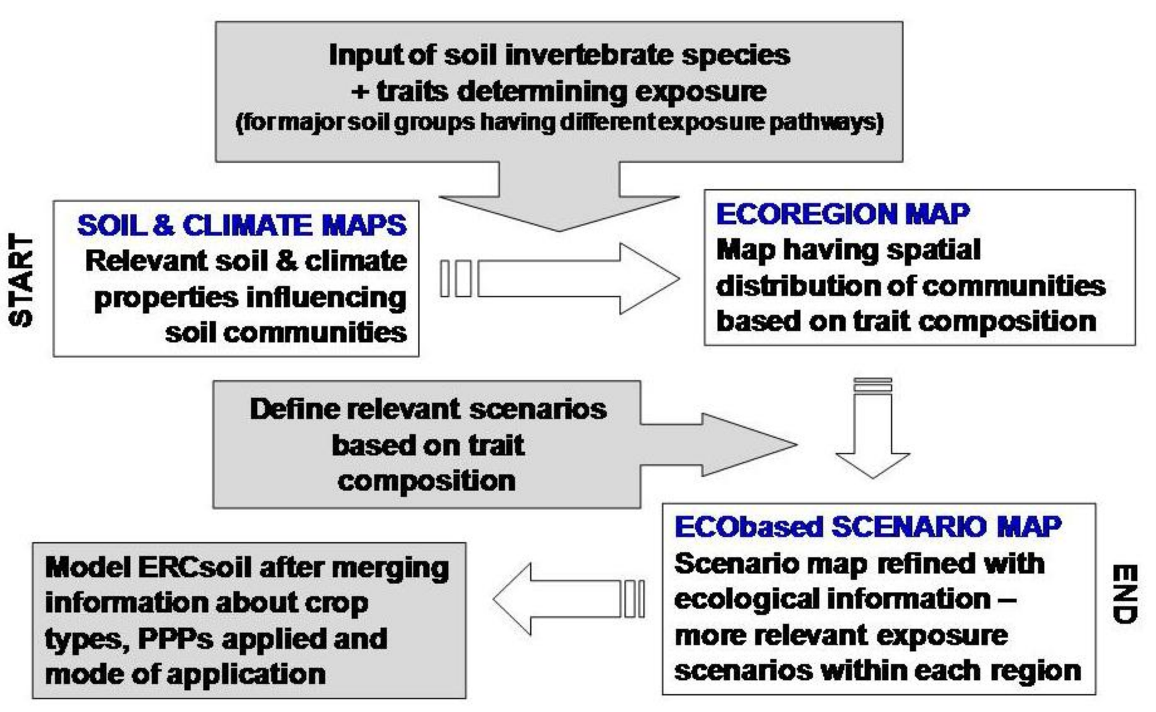


Figure 5: Flow-chart of the derivation of ecoregions in the EU

Table 1: Soil depth profiles where the life form groups are exposed to pesticides. The litter layer is considered particularly relevant for permanent crops or minimal tillage crops (for more details see EFSA, 2010b). Note that these soil depth profiles, with the exception of the litter layer, are currently being considered in the work related to the update of the persistence in soil guidance document (EFSA, 2010a).

	Depth profile where the organisms are exposed					
	Litter layer	**0 – 1 cm**	**0 - 2.5 cm**	**0 – 5 cm**	**0 – 20 cm**	**burrows**
Enchytraeids	litter dweller	litter dweller	intermediate	mineral dweller		
Earthworms	epigeic + anecic	epigeic + anecic			endogeic	anecic
Isopoda	litter dweller	litter dweller		soil dweller		
Collembola	epigeic	Epigeic	hemiedaphic	euedaphic		

To illustrate this concept two examples are given:

Example 1 : Two ecoregions with similar set of exposure profiles: Potential soil communities with similar "trait groups" (see also Figure 6)

– In an acid sandy soil in **Portugal** the community is dominated by mesofauna (Collembola) and isopods, accompanied by a low abundance of endogeic earthworms. Thus depending on the crop type, the exposure in the top 2.5 cm of soil (or top 5 cm depending on the collembolan trait groups) and the litter layer should be modelled.

– In an acid sandy soil in **Germany** the community is also dominated by mesofauna (Collembola) and diplopods, accompanied by a low abundance of endogeic earthworms. So, the recommendation regarding which exposure should be modelled is equal to Portugal.

Example 2: Two ecoregions with different set of exposure profiles: soil communities with a different set of "trait groups" (see also Figure 7)

- In a clay soil in southern **Portugal** the community is dominated by mesofauna (Collembola), accompanied by a low abundance of endogeic earthworms, meaning that modelling the exposure situation in the top 5 cm of soil is sufficient.
- In a clay soil in **Germany** there are many earthworms (endogeic and deep burrowing anecics). Here the exposure in the top 20 cm of soil as well as in burrow linings should be modelled.

It should be noted that the limit below which a group of organisms is to be neglected in terms of exposure modelling ("low abundance of organisms") has to be considered case by case by risk management authorities.

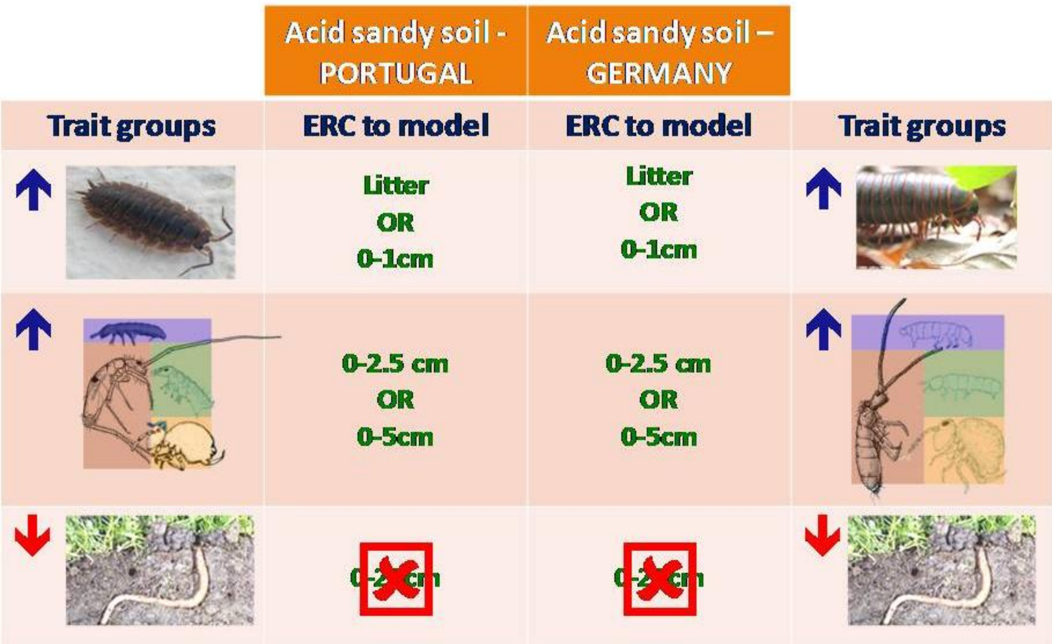


Figure 6: Example of two Ecoregions with a similar set of exposure profiles. Arrows represent the level of dominance: arrow up high dominance, arrow down low dominance.

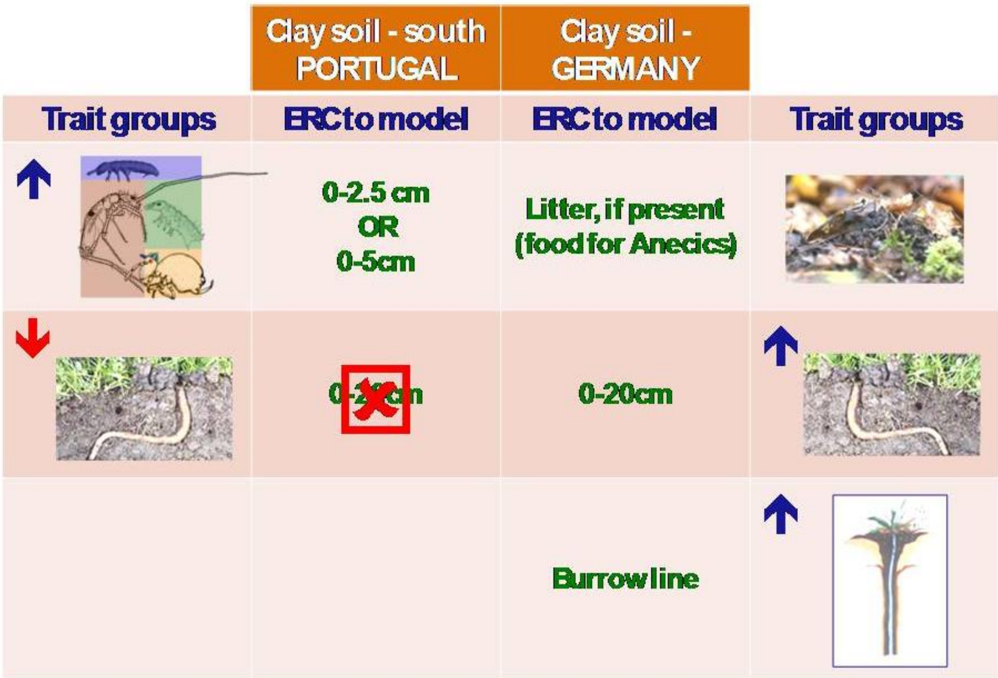


Figure 7: Example of two Ecoregions with different sets of exposure profiles. Arrows represent the level of dominance: arrow up high dominance, arrow down low dominance.

4. Testing the concept

Defining an ecoregion map focusing on soil organism communities, and developing region-specific ecologically relevant exposure scenarios, implies the collation of different types of information. Besides the compilation of environmental data such as soil, land-use and climate on a geographical basis, it is of paramount importance to collect ecological and geographical distribution data for soil fauna. This will allow us to define their relative importance within each ecoregion and to define the relevant soil layers where organisms are exposed.

For this purpose a database containing biogeographical information on selected key soil fauna groups with a relevant ecological role in European soils was compiled. The fauna groups selected for the database represent different morphological and ecological characteristics influencing exposure. Data on the presence of earthworm, enchytraeid, collembolan and isopod species were collected from the literature for three countries representing distinct biogeographical regions in Europe: Finland, Germany and Portugal. Criteria used to select the organism groups and the model countries, as well as the structure of the compiled data base are presented in the next sub-sections.

4.1. Selection criteria for the model invertebrate groups

Due to the fact that the soil community is very diverse and the richness of soil organisms in a certain location can easily exceed several hundreds of species (e.g. in beech wood forest in Southern Germany; Beck et al., 1988), it is simply impossible to compile biogeographical information for every soil group. Therefore, a small number of organism groups representing different functional roles in soils were selected by the PPR Panel to testing the concept.

The following selection criteria were used, listed in order of importance:

1. Important ecological role in European soils, e.g. in terms of biomass, soil structuring activity, and place in the food web.
2. Presence across a wide geographical scale, e.g. Northern European and Mediterranean soils should be covered.
3. Different morphological and ecological characteristics influencing exposure, e.g. different size classes (meso- as well as macrofauna), soft-bodied versus hard-bodied groups (oligochaete worms vs. arthropods mainly insects).
4. Availability of information regarding their distribution, preferably in data bases, maps or at least review papers.
5. Availability of trait data on the selected groups (preferably in data bases), particularly life-form traits indicating at which soil depth they are mainly active.
6. Groups including species being regularly used in ecotoxicological testing. This would help at a later stage when combining information from exposure modelling and effect testing.

The following combination of four groups fulfils these criteria:

1. Collembola (springtails): Mesofauna, hard-bodied, important microbial regulators during the decomposition process, widely distributed with many species all over Europe;
2. Isopoda (woodlice): Macrofauna, hard-bodied, most species prefer warmer regions; important detritivores in the early stages of organic matter decomposition (usually called "litter transformers")
3. Lumbricidae (earthworms): Macrofauna, soft-bodied, important microbial regulators often with very high biomass, key group for soil structure formation and maintenance, widely distributed in Europe.
4. Enchytraeidae (potworms): Mesofauna, soft-bodied, important microbial regulators often in very high numbers, prefer cool, acid soils.

This selection of groups fits with recommendations recently made for biological soil monitoring in the EU. Sampling of earthworms (plus enchytraeids), springtails and soil micro-organisms was recommended by the EU funded FP6 ENVASSO[10] project for a first tier, while other organism groups (e.g. nematodes) could be used to address specific biodiversity monitoring questions (Bispo et al., 2009).

The criteria for the selection of biodiversity indicators adopted by ENVASSO have ecological relevance as the utmost condition for selecting an organism group. Based on this criterion other soil fauna groups could have been selected. Nematodes are one example, not only due to their ecological relevance, but also to the well established functional classification. However the existing biogeographical information (key criterion also in this case) is scarce and limited to a few countries within the EU (e.g., The Netherlands), which limits their use in this case. This kind of limitation is also present for other soil invertebrate groups like e.g. soil mites, diplopods and slugs. For micro-organisms, despite their dominance and fundamental relevance for the processes in soil, until now field distribution data has been scarce, although very recently some progress has been made (Fierer and Jackson, 2006; Lauber et al., 2009; Lauber et al., 2008; Ranjard et al., 2010).

[10] ENVASSO: ENVironmental ASsessment of Soil for mOnitoring, EU FP6, Contract No: 022713

4.2. Selection criteria for the model countries

The selection of the model countries was based on the coverage of different biogeographical regions in Europe. The selection of the countries should attempt to maximise the differences in climate and soil properties, thus hosting different soil organism communities. Also the availability of data for the selected soil organism groups (in published papers and/or in databases) was considered for the selection.

The following assumptions were made:

- Ecoregions differ with respect to presence and abundance of characteristic species of the selected taxa;
- Ecoregions differ with respect to the relative number of species in different life forms within one taxon;
- Important ecological functions (e.g. organic matter breakdown) in different ecoregions are carried out by different taxa;
- The vertical distribution of life forms differs between ecoregions.

Based on these criteria, the model countries selected were Finland (representative of the Boreal region), Germany (representative of the Continental region), and Portugal (representative of the Mediterranean region), to achieve the best coverage of the North-South gradient in Europe.

4.3. Construction of the database

4.3.1. Biogeographical data

For each organism group considered, one database per country was built. Each data base had a similar structure composed of four sections as follows (for more details see Appendix A):

1. Section 1 – Site information, containing data on site location (geographical coordinates), land-use type and, when available, dominant vegetation;
2. Section 2 – Soil type information, containing data on major soil properties;
3. Section 3 – Species information, containing the taxonomic data for each species, abundance or density (when available), and the sampling method used to collect the data. In this section the information related to the life-form type (see a-d below) is also included. This was defined using morphological and ecological traits available in published material and/or in trait databases. The following life-form groups were defined:
    a. Collembola[11]
        i. Euedaphic species with very low dispersal ability, living down to 5 cm[12]
        ii. Hemiedaphic (medium dispersal) species, living down to 2.5 cm
        iii. Epigeic (fast dispersal) species, living at the soil surface
    b. Earthworms
        i. Anecic species that live in permanent vertical burrows in mineral soil layers up to 3 m deep and feed from the litter layer

[11] Based on the 3 morphological traits, five life-form classes were defined in the database. However, for the purpose of this opinion (and for data analysis) these were grouped into the 3 classes described.
[12] In ploughed soil, these organisms can be distributed over the whole ploughing horizon, e.g. up to 20 cm.

   ii. Endogeic species that inhabit mineral soil, making horizontal non-permanent burrows, mainly in the uppermost 20 cm of soil

   iii. Epigeic species that live above the mineral soil surface, typically in the litter layers of forest soils, making no burrows

c. Enchytraeids:

   i. Soil dwellers: species that live mainly in soil down to 5 cm depth

   ii. Intermediate dwellers: species that circulate between soil and litter layers

   iii. Litter dwellers: species that live mainly in the litter layer

d. Isopods

   i. Litter dwellers: species living mainly on the soil surface, particularly in the litter layer

   ii. Soil dwellers: species living mostly in the soil surface, but that are able to burrow down to 2.5 cm depth

4. Section 4 – Bibliographic references, containing the complete information for all references included in the database.

Over 200 references were surveyed and data from 168 papers and reports were introduced in the database (for more details refer to Appendix A). Duplicate references, i.e., different scientific papers reporting the same type of taxonomic information for the same sites, were not considered. Overall data compilation resulted in extensive databases whose information is summarized in Tables 2 and 3.

Regarding the completeness of published material analysed, for Portugal, with the exception of enchytraeids, a good coverage was achieved. The lack of biogeographical data observed for enchytraeids is due to the fact that, for this group, only one study was reported. For Germany a good coverage was also achieved for collembolans, earthworms and enchytraeids. However, results of soil biodiversity monitoring activities from several Federal States may not have been considered due to the fact that they are not published (Gardi et al., 2009). For isopods, the difficulty in obtaining old taxonomic papers made it difficult to cover all species reported for Germany and, most importantly, to have a good spatial distribution of data (see next point). For Finland, a generally good coverage of the published literature was achieved for all organism groups.

It should be noted that many data sets were not included due to doubts on their quality. For example, when the geographical information was not sufficient to clearly identify the sampling site, the data sets were not included. In the case of enchytraeids, references to the genus *Friderica* (the most species-rich genus in this family) were not included if they were not confirmed by a specialist who recently reviewed this genus.

Despite the overall good geographical coverage, it was not possible to complete all information for every data entry in the database. Several constraints were found, namely related to the nature of papers analysed. Many of those were taxonomic papers containing no precise information on the geographical location or on the land-use where the biological material was collected. Also, information on the abundance or the soil properties was often not available.

One aspect clearly visible from this data overview is the low number of entries related to crop areas. The reason is that most ecological and taxonomic studies were performed in forested and grassland areas. However, this fact is not expected to affect the definition of ecoregions to develop ecologically relevant exposure scenarios to plant protection products. The use of data from non-crop sites is acceptable due to the following reasons:

- Soil communities in agricultural soils are a relic from forest soil communities. They have fewer species from the same species pool.
- For each organism group considered, the community structure in terms of life-forms is similar between arable and other land-uses within the same geographical area (Table 4).

Moreover, the environmental risk assessment of plant protection products may need to be done for in-crop and off-crop areas, and also at a broader spatial scale (see EFSA 2010c, 2010d). In this context forest areas, especially those integrated in the agricultural landscape, acquire an extra importance because:

- they constitute a genetic reservoir for biodiversity at landscape and regional level
- they can act as donor areas for external recovery of communities in crop areas

Table 2: Number of data entries/sites covered per land-use type and per organism group in each country. Numbers in brackets indicate no information on the land use type. Empty cells imply that no records were available for this land use.

Organism group	FINLAND			
	Forests	Grasslands[a]	Arable crops	Others
Collembola	1052/16	27/2	212/3	20 (839)/1(27)
Isopods			3/2	(185)/(69)
Enchytraeids	164/15			
Earthworms	96/10	49/7	5/3	8/3
Organism group	**GERMANY**			
	Forests	Grasslands[a]	Arable crops	Others
Collembola	702/44	215/14	500/15	255 (31)/21 (14)
Isopods	35/11	20/8	2/2	1 (28)/1 (28)
Enchytraeids	611/50	489/51	38/9	
Earthworms	338/54	621/73	147/20	(15)/(1)
Organism group	**PORTUGAL**			
	Forests	Grasslands[a] / Shrublands	Arable crops	Others
Collembola	2129/51	179/5	91/5	474 (3)/38 (3)
Isopoda	47/4	49/6	59/7	8 (97)/2 (10)
Enchytraeids			7/1	
Earthworms	21/9	20/5	6/4	7(15)/4 (11)

[a] meadows and pastures are included in grasslands

Table 3: Number of species per land-use type and per organism group in each country. Sum is usually higher than total number of species since several species are present in more than one land-use type. Numbers in brackets indicate no information on the land use type. Empty cells imply that no records were available for this land use.

Organism group	**FINLAND**				
	Forests	**Grasslands[a]**	**Arable crops**	**Others**	**TOTAL**
Collembola	70	18	44	19 (199)	220
Isopods			3	(25)	26
Enchytraeids	25				25
Earthworms	14	12	4	2	19
Organism group	**GERMANY**				
	Forests	**Grasslands[a]**	**Arable crops**	**Others**	**TOTAL**
Collembola	214	112	116	146 (20)	334
Isopods	12	7	2	1 (28)	43
Enchytraeids	66	62	19		87
Earthworms	40	35	22	(15)	54
Organism group	**PORTUGAL**				
	Forests	**Grasslands[a]/ Shrublands**	**Arable crops**	**Others**	**TOTAL**
Collembola	245	96	58	158 (3)	303
Isopods	45	47	25	4 (70)	115
Enchytraeids			7		7
Earthworms	13	15	6	5(13)	36

[a] meadows and pastures are included in grasslands

Table 4: Average number of species per life-form class for each organism group on the main land-use types in each country (sites with only one species identified were discarded from this analysis). Empty cells imply that no records were available for this land use.

Life-form class		FINLAND			GERMANY			PORTUGAL		
		Arable crops (10 sites)	**Forest (12 sites)**	**Grassland (no sites)**	**Arable crops (12 sites)**	**Forest (24 sites)**	**Grassland (7 sites)**	**Arable crops (18 sites)**	**Forest (54 sites)**	**Shrub areas (4 sites)**
Collembola	Euedaphic	3	4		5	5	5	2	3	6
	Hemiedaphic	6	8		11	8	7	10	10	18
	Epigeic	7	9		13	11	10	9	9	19
		Arable crops (1 site)	**Forest (8 sites)**	**Grassland (4 sites)**	**Arable crops (28 sites)**	**Forest (47 sites)**	**Grassland (60 sites)**	**Arable crops (1 site)**	**Forest (4 sites)**	**Grassland (3 sites)**
Earthworms	Anecic	0	1	1	1	0	1	0	0	1
	Endogeic	1	2	3	3	2	3	2	2	2
	Epigeic	2	4	4	1	3	1	1	2	4
		Arable crops (no sites)	**Forest (13 sites)**	**Grassland (no sites)**	**Arable crops (12 sites)**	**Forest (46 sites)**	**Grassland (37 sites)**	**Arable crops (1 site)**	**Forest (no sites)**	**Grassland (no sites)**
Enchytraeids	Litter dweller		4		0	2	1	1		
	Intermediate		2		1	2	1	3		
	Soil dweller		3		3	3	5	3		
		Arable crops (1 site)	**Forest (no sites)**	**Grassland (no sites)**	**Arable crops (no site)**	**Forest (4 sites)**	**Grassland (6 sites)**	**Arable crops (4 sites)**	**Forest (3 sites)**	**Shrub areas (5 sites)**
Isopods	Soil dweller	2				1	0	3	2	1
	Litter dweller	0				5	3	12	14	8

4.3.2. Land use, soil and climate data

The database of soil organism observations (biogeographical database), which was available for several points in Finland, Germany and Portugal, was converted into a geographical database, associating geographical coordinates to each observation site. This process meant that it was then possible to locate the survey points over a geographic map of Europe.

The biogeographical database consists of data on presence/absence, and in some case abundance, of selected soil organism groups. For some entries data on land use, vegetation, soil and climate were also reported. These environmental parameters, essential for the ecological characterization of soil community, were however very incomplete (see section 3.4). For this reason, alternative data on land use, soil and climate were used to fill-in the gaps present in the original dataset. The following data sets, provided by JRC, were used in the process. For more details, please refer to Appendices B and C.

Grid maps for the following parameters:

a) Organic matter in soil (kg/kg) is derived from the Organic Carbon Topsoil map.
b) Soil Texture to derive the soil and water content is available in the SGDBE[13] at the level of soil mapping units. It was converted from vector to grid cells.
c) Soil pH
d) Land-use from Corine Land Cover 2000
e) Temperatures (mean, maximum and mininum)
f) Mean annual precipitation

All data have been provided in a common resolution and with common projections (e.g. INSPIRE[14], reference grid 10x10 km^2 to be used in the SGDBE v2), with a resolution of 1x1 km^2.

This process has been carried out using the utilities of spatial analysis present in a Geographical Information System (GIS). Once the geographic position of a sampling point is known, it is possible to do a spatial query in the GIS, concerning the values of soil pH, organic matter, total precipitation, and any other parameter that is available in the form of a geographic database (Figure 8).

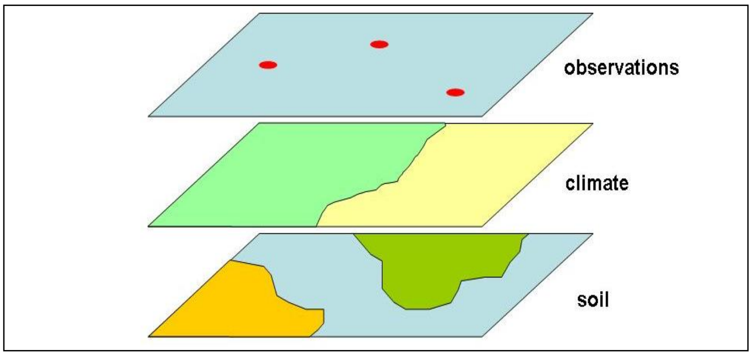


Figure 8: Schematic representation of the spatial query used for the construction of the geographical database.

[13] SGDBE: Soil Geographic Data Base of Europe

[14] INSPIRE: Infrastructure for Spatial Information in Europe; The INSPIRE directive came into force on 15 May 2007 and will be implemented in various stages, with full implementation required by 2019. The INSPIRE directive aims to create a European Union (EU) spatial data infrastructure.

4.4. Data Analysis Methodology

The methodology used for the data analysis is described in detail in a separate scientific report (EFSA, in press). In this section a summary of the methodology is presented. For more details please refer to the above mentioned report.

4.4.1. Data Structure

The main task of the data analysis was the description of the regional distribution of different trait groups of selected soil invertebrate groups in Europe, depending on site climatic and soil characteristics. Hereby several restrictions in the biogeographical database limited the possibilities for data analysis:

- incomplete information on site locations,
- missing information on climatic descriptors,
- high percentage of missing values of soil parameters,
- missing information on species abundance.

To fill-in the data gaps, originated by the lack of information in the consulted literature, surrogate data on site-specific climatic and soil characteristics from different databases were used, which are described in more detail in section 3.5. These data describe Europe on a 1 km² scale and were linked through the site UTM (Universal Transverse of Mercator) coordinates to the biogeographical database. Missing coordinates in the biogeographical database were filled-in by deriving coordinates from the given name of the site (region, village/town, name/place, and additional site info). For 2.7% (266) of the records in the biogeographical database no specific site information was given, e.g. only the country or the region was given. These data were excluded from the analysis. For most records the local scale of biological sampling is much smaller than the extent of the site. Therefore, the records in the biogeographical database which are corresponding to the same set of UTM coordinates and equal land use were assumed to originate from one site. This sometimes combines samples from several locations within one site. Some UTM coordinates specify a grid cell in the soil database, which is classified as “urban area”. For these grid cells, no soil parameters were available. More information on the biogeographical database and the comparison between literature and surrogate data can be found in the technical report (EFSA, in press).

The following information was used from the biogeographical database:

Name of the country: Location of the site in Finland, Germany or Portugal.

Information on land use: The given information was classified into 14 categories of the CORINE system (EEA, 2000): 131 (Mine, Dump and Construction sites), 141 (Artificial, non-agricultural vegetated area), 211 (Arable land), 231 (Pastures), 242 (Complex cultivation pattern), 243 (Land with significant natural area), 244 (Agro-forestry areas), 313 (forests, including 311, 312 and 313 for the purpose of this data analysis), 323 (Sclerophyllous[15] vegetation), 331 (Open spaces with little or no vegetation), 411 (Inland wetlands), 421

[15] Hard leaf vegetation definition from CORINE (EEA, 2000): “Bushy sclerophyllous vegetation, includes maquis and garrigue. In case of shrub vegetation areas composed of sclerophyllous species such as Juniperus oxycedrus and heathland species such as Buxus spp. or Ostrya carpinifolia with no visible dominance (each species occupies about 50% of the area), priority will be given to sclerophyllous vegetation and the whole area will be assigned class 323.”

(Marine wetlands), 998 (Cave) and 999 (Greenhouse)[16]. In the data analysis some CORINE classes were pooled as follows: “Crop land” (211, 242, together 11%), “Grass land” (231, 13%), “Forest” (244, 313, together 43%) and “others” (131 (0.2%), 141 (3.3%), 243 (2.3%), 323 (0.9%), 331 (1.9%), 411 (2.6%), 421 (1.1%), 998 (1.5%), 999 (0.04%)), which were rare in the database.

Occurrence of a specific species: All information on species was checked for consistency and unique naming. A species was counted as present on a site when recorded at least once.

Classification of the life form group: The life form group of each species was classified according to the typology described in section 4.3.1.

The following information on the soil was added from the JRC database (see Section 4.3.2 and Appendix B):

The pH-value was used as in the database.

The Organic Carbon content was estimated from the organic matter content by a linear pedotransfer function (factor 1/1.724) (FOCUS, 2000).

The Sand, Silt and Clay content was estimated from the mean values of 6 classes (coarse, medium, medium fine, fine, very fine and full organic) (see for details EFSA, in press).

Table 5: Conversion of soil texture classes (JRC) to sand silt and clay content.

JRC Code	Description	Clay [%]	Silt [%]	Sand [%]
9	Full organic	0.0	0.0	0.0
5	Very fine	73.3	13.3	13.3
4	Fine	46.5	26.7	26.7
3	Medium fine	17.5	75.0	7.5
2	Medium	18.0	39.4	42.6
1	Coarse	7.6	13.7	78.7

The following information on the climate was added from the JRC database (see Section 4.3.2. and Appendix B):

The total annual Precipitation was used as in the database

The annual Mean Temperature was used as in the database

The Range of the Temperature was estimated by the difference of maximum and minimum average monthly temperature within 1960-1990.

The potential Evapotranspiration was used as in the database

[16] Samples for the inland wetlands (2.61 %), marine wetlands (1.06 %), caves (1.51 %) were always from borderline situations (e.g. they were taken on a soil spot or at the cave entrances). The data for greenhouses (0,04%) can be even considered as having a negligible effect on the results

For each site (combination of UTM coordinates and land use) the number of different species per life form group was counted. The percentage of a life form group in relation to the total number of species defines the **raw relative richness** of that life form group on this site.

4.4.2. Calculation of adjusted relative richness

The potential maximum number of species for each life form group within a specific taxon can be very variable. For example, in earthworms the potential maximum of anecic species is 2, while for endogeic or epigeic species it is 7 (see Table 6). Thus the raw relative richness of less rich life form groups could be under-represented. In order to compensate for this bias, for each organism group the absolute number of species per trait group was divided by their maximum and limited to 100%, giving the **relative occurrence of species of a specific life form group** per site. The percentage of the relative occurrence in relation to the sum of all trait groups defines the **adjusted relative richness of that life form group** on this site (Table 6). This variable was used as dependent variable of the analysis in three of the four groups (earthworms, enchytraeids and isopods). However, for Collembola the overall number of species is higher than for earthworms and enchytraeids and, additionally, their species distribution over the 3 life form groups is more homogeneous. Therefore, the analysis was done with the raw relative richness.

The definition of the maximum number of species per life form group within each taxon was defined by expert judgement, especially based on the knowledge of several datasets from different areas in Europe. The empirical verification showed that only in a few records were these maximum numbers exceeded. For instance, in earthworms, this was mostly due to allochthonous species at the respective sites as for example, species introduced by human activities.

The maximum numbers of species and the relative richness measures used for modelling (and derived from the models) for each organism group are shown in Table 7.

For comparing the distribution of sites, a triangular graph was used (Figure 9)[17]. With this approach, the composition of soil animal species assigned to three life form types can be visualised in an easily comprehensible manner. For those organism groups where the adjusted relative richness was calculated, the distribution of values, is more balanced in the centre of the triangular graph and displays geographical patterns better than the raw relative richness (Figure 10 exemplifies this for earthworms).

Table 6: Example for the calculation of the Adjusted Relative Richness

Earthworms at a specific site (combination of UTM coordinates and land use; see Section 3.4.1. for details)				
Life form group	**anecic**	**endogeic**	**epigeic**	Sum
Number of different species	2	2	4	8
Raw Relative Richness	**25%**	**25%**	**50%**	100%
Maximum number of species	2	7	7	14
Relative occurrence	100% (=2 of 2)	29% (=2 of 7)	57% (=4 of 7)	186%[18]
Adjusted Relative Richness	**54%**	**15%**	**31%**	100%

[17] The presentation of data referring to various percentages of three factors which are summed up to 100% is frequently used for the derivation of soil texture classes in pedology (soil texture is defined by the percentage of sand, silt and clay particles; Ad-hoc AG Boden, 2005; Blume et al, 2010).

[18] This value was used in the calculations but it is not related to any biological meaning.

Table 7: Maximum number of species per life-form group relative richness measure used in the modelling approach (and respective output variable) for the four soil organism groups considered.

Earthworms	**Anecic**	**Endogeic**	**Epigeic**
Maximum nº of species	2	7	7
Variable used in model	Observed adjusted relative richness		
Model output	Fitted adjusted relative richness		
Enchytraeids	**Soil dwellers**	**Intermediate**	**Litter dwellers**
Maximum nº of species	6	6	6
Variable used in model	Observed adjusted relative richness		
Model output	Fitted adjusted relative richness		
Collembola	**Euedaphic**	**Hemiedaphic**	**Epigeic**
Maximum nº of species	10	16	24
Variable used in model	Observed raw relative richness		
Model output	Fitted raw relative richness		
Isopods	**Soil dwellers**	**Litter dwellers**	
Maximum nº of species	6	10	
Variable used in model	Observed adjusted relative richness		
Model output	Fitted adjusted relative richness		

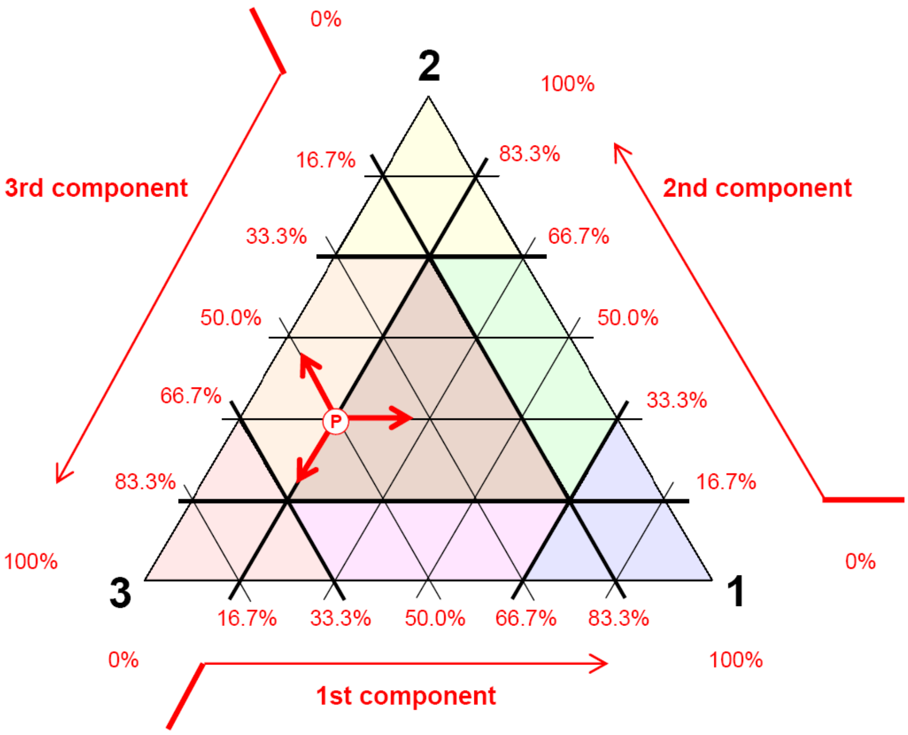


Figure 9: Triangular graph to visualize a composition of three components, e.g. life forms. The point (p) indicates a composition of 16.7% of the first component, 33.3% of the second and 50% of the third. The coloured zones indicate dominance classes according to Figures 12 and 13.

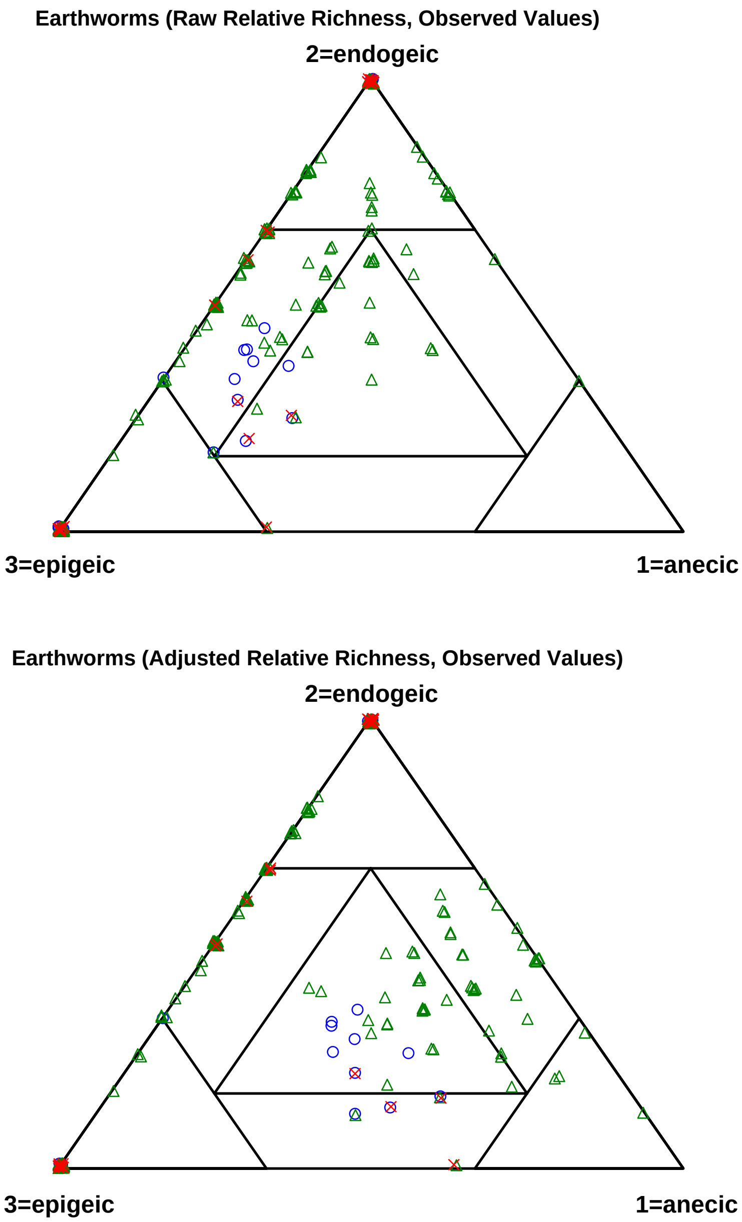


Figure 10: Triangular graphs of observed, relative richness ***(top)*** and of observed, adjusted relative richness ***(bottom)*** of different life forms of earthworms for Finland (blue circles), Germany (green triangles), Portugal (red crosses). For scale and categories see Figures 13 and 14.

4.4.3. Modelling

To analyse the dependence of the adjusted relative richness (or raw relative richness in case of Collembola) to the site characteristics, the adjusted relative richness was further transformed to a 2-dimensional parameterisation of the triangle using the first component and the conditional value of the second component given that the first does not appear. Both values can independently range from 0 to 100% and were modelled using a generalized linear model with logit link function, binomial variance and additional scale parameter (σ^2). The computations were done with the SAS GENMOD procedure (SAS, 2004). For more details please refer to EFSA (in press).

As independent variables, the three indicators for land use ("crop land", "grass land", "forest"), the average pH value, the average content of organic carbon, clay and sand, the total precipitation, the mean temperature and the range of temperature was used per site. For the fit of collembolan and isopod data additional information on evapotranspiration was used. To avoid overfitting no higher orders or interactions were introduced.

The complete model was further simplified using a stepwise backwards selection strategy. An independent variable was excluded when the sum of the p-values (single factor likelihood ratio test (SAS, 2004)) of both components was highest and both components had a p-value above 0.2.

Table 8: Independent site characteristics selected to model the adjusted relative richness of organism groups

Group of organism	*Selected independent variables*
Earthworms	Land use indicator, pH value, organic carbon content, precipitation, temperature, range of temperature
Enchytraeids	pH value, organic carbon and clay content, temperature, range of temperature
Collembola	All variables (no selection strategy applied)
Isopods	Clay and sand content, precipitation and evapotranspiration

The empirical data show an overdispersion of about 30%, which can be interpreted as additional variation of the data, which is not explained by the model. A complete description of the data analysis can be found in a separate scientific report of EFSA (in press).

This model allows the calculation of the **fitted adjusted relative richness** (or the **fitted raw relative richness** in the case of Collembola) for each site with given characteristics. Figure 10 shows the results for the given sites of the biogeographical database for earthworms. The model clearly tends to push the data from the borders to the centre and to cluster the data. There are two reasons for this result. First the logit model does not allow extreme relative richness values with zero or 100% components. This reflects the idea that the model describes the **potential** distribution of the life form groups and not their **actual** occurrence in the sample. We assume that, on each site, there is at least a small chance for each trait group to occur. The second reason is the difference in scale between the observations and the independent variables of the model. While the observations mainly reflect the situation on a specific local spot with several additional influencing factors, the model describes the relative richness on average parameters of the underlying 1 km² grid. The fitted values should therefore reflect the general tendency of the location and not the properties of the specific site. In consequence, the differentiation by country appears clearer in the triangle with modelled (fitted) relative richness (Figure 11) than with the observed values (observed adjusted relative richness – Figure 10).

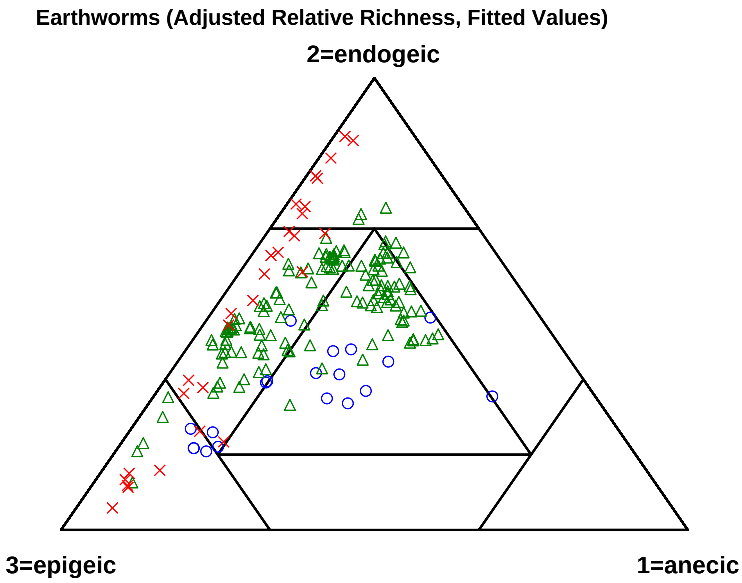


Figure 11: Triangular graph of **modelled, adjusted** relative richness of different life forms of earthworms / selected variables. Finland (blue circles), Germany (green triangles), Portugal (red crosses). For scale and categories see Figure 13.

To transfer these data to the European scale and to draw a European ecoregion map, an additional simplification has to be applied. For this purpose, dominance classes of life form groups were defined. We used a simple geometrical approach to divide the triangle into dominance classes (Figures 12 and 13). The central 25% of the area are categorized as equally distributed, with no particularly dominating life form group. Two life form groups are regarded as dominating the distribution when the sum of their relative richness (adjusted or raw, depending on the organism group considered) is equal to or higher than 83.3%. If not, all three life forms are considered important. Furthermore a single life form group is dominating when its relative richness is higher than 66.7%. (Figures 12, 13 and 14).

Using these rules, the **potential adjusted (or raw) relative richness** of a group of species can be calculated for each grid cell of the European map, classified in 7 dominance classes and drawn in a map (Figure 15).

The comparison of observed (colours on the dots) and fitted values on the map (Figure 15) shows the modelling effects explained before. Since local variation is not fully covered by the map, extreme observations (blue, yellow, red) are smoothed to more homogeneous distributions. This model effect can also be seen in Table 9 showing the concordance between the observed and modelled number of sites for each life form dominance classes (here exemplified for earthworms).

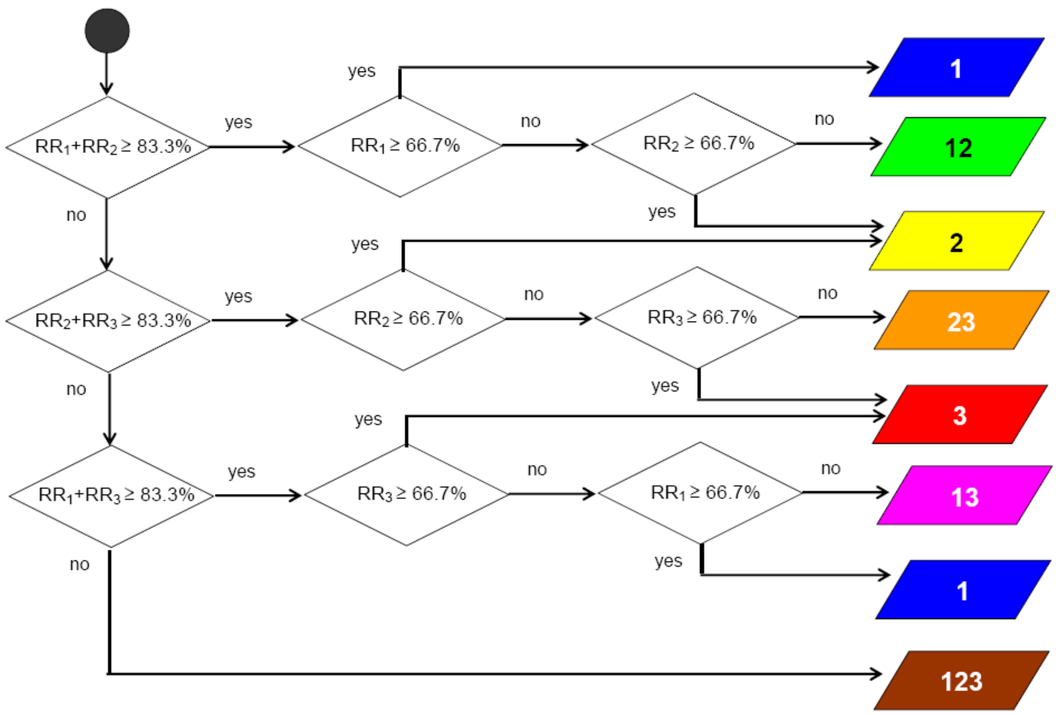


Figure 12: Example of categorization rule of the relative richness (RR) into dominance classes of three different life forms for earthworms (called 1, 2, and 3 in this graph) or their respective combinations (12, 23, 13, and 123).

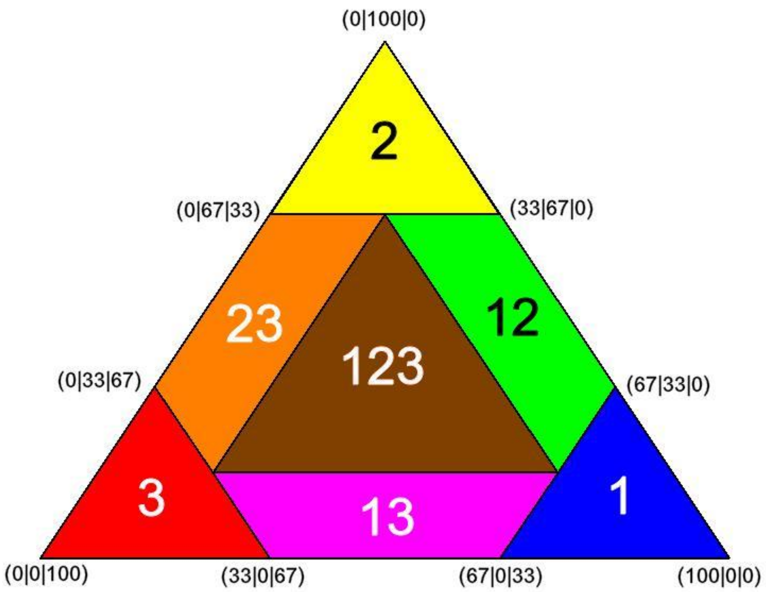


Figure 13: Example of categorization into dominance classes of the relative richness (adjusted or raw, depending on the organism group considered) of three different life forms for earthworms (called 1, 2, and 3 in this graph) or their respective combinations (12, 23, 13, and 123). Coordinates (e.g. 0/33/67) are given in percent.

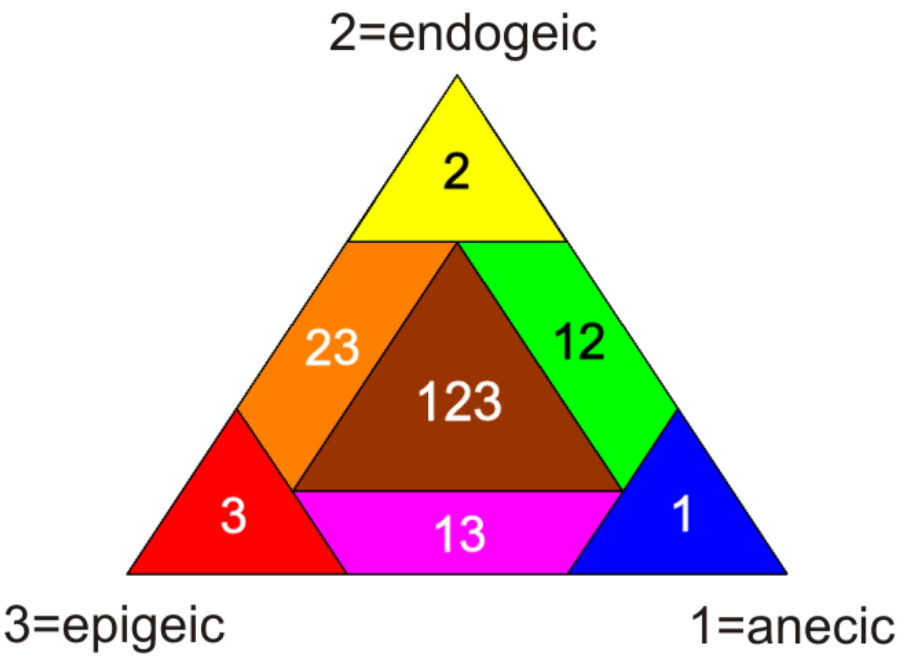


Figure 14: Categorisation of the adjusted relative richness of three different life forms of earthworms into dominance classes (called 1, 2, and 3 in this graph) or their respective combinations (12, 23, 13, and 123).

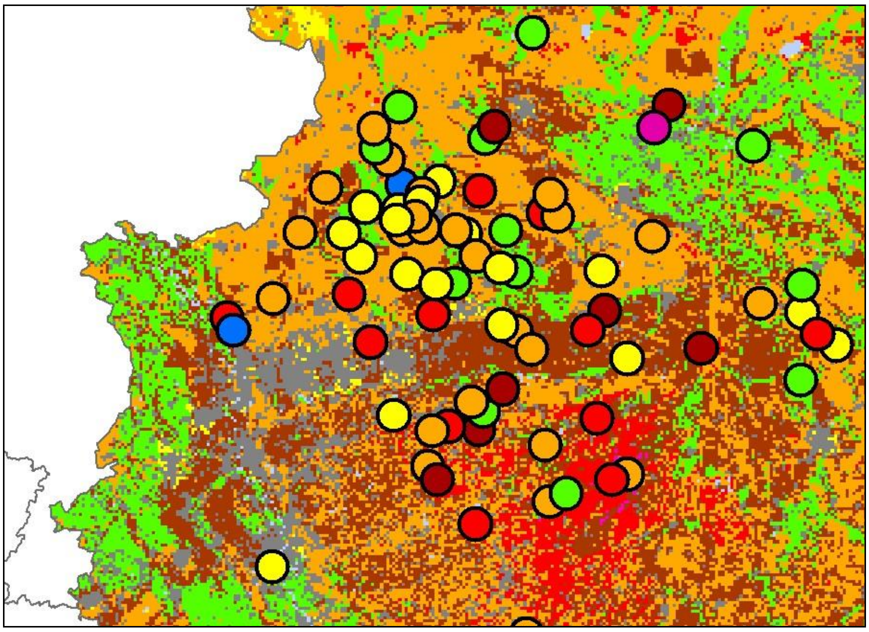


Figure 15: Dominance classes of earthworms in the German "Ruhrgebiet". Colours according to the classification scheme (grey = urban area). Single dots show the observations with their observed adjusted relative richness.

Table 9: Concordance of observed and modelled life form dominance classes for adjusted relative richness of earthworms (202 sites)

No of Sites		Modelled Life Form Category							
		1	**2**	**3**	**12**	**13**	**23**	**123**	Sum
Observed Life Form Category	**1**	0	0	0	0	0	1	3	4
	2	0	7	1	0	0	25	12	45
	3	0	1	8	1	0	22	6	38
	12	0	2	0	1	0	5	17	25
	13	0	0	2	0	0	4	1	7
	23	0	1	3	1	0	26	11	42
	123	0	0	3	2	0	9	16	30
	Sum	0	11	17	5	0	92	66	191(a)

(a): For 11 sites no fit was possible

4.5. Soil Ecoregion Mapping

From the data analysis on the complete biogeographical database and the associated environmental variables, a provisional model was obtained, according to the procedures described in section 4.4.3. The model was then applied, using JRC data as input parameters, in order to estimate the distribution of earthworms, collembola and enchytraeid dominance classes of life form groups, in the three investigated countries. All the spatial analyses, including the extraction of soil and climate parameters in selected sites, the spatial query and the spatial overlay, have been performed using ArcGIS 9.3.1. A more accurate description of these procedures is provided in Annexes B and C.

The output of these models was a series of maps (one for each organism group and country). The territories of Finland, Germany and Portugal were classified in seven classes, according to the invertebrate data corresponding to these countries as shown in the triangles reported in Figures 12 and 13.

Earthworm ecoregion maps were produced for the three investigated countries, but restricting Finland to its Southern part because no sampling data in the Northern part were available. Enchytraeid ecoregion maps were limited to Germany and Finland since almost no enchytraeid data were available for Portugal (1 site only). Due to the incompleteness of soil data sets in urban areas, the Ecoregion maps have been masked in these areas. No maps were produced for Collembola and Isopoda (for details refer to Section 4.6.1.3 and 4.6.1.4).

Although in principle the interpolation over the entire EU territory would have been technically feasible, mapping of territories without observed values was not considered to be reliable for the purpose of this opinion.

The concepts of exposure scenario and the definition of soil profile depth relevant for different soil organism communities led to the production of profile maps for earthworms and enchytraeids. The territory of the investigated countries was classified on the base of the soil depth relevant for risk assessment. The depths on the profile map were specifically assigned to the preferred microhabitat of each organism group, e.g., litter layer, soil surface (0-1 cm), 0-2.5 cm, 0-5 cm or 0-20 cm. This profile was further refined to define a "worst case" profile (Figure 16, see also Figures 18 and 28).

For earthworms, each trait group could be associated with typical soil depths where the organisms are living and thus exposed: anecic earthworms are mainly exposed via food consumption at the soil surface (litter or 1 cm)[19] and - via contact - in their burrows, epigeic earthworms are also exposed in the soil surface (Table 10). On the contrary, endogeic earthworms are exposed in a soil depth of 20 cm. This "translation" leads to a reduction of complexity of the maps, reducing the 7 combinations of trait groups to 3 possible combinations of depth profiles: soil surface (litter or 1 cm), 20 cm, and the combination of both (litter or 1 cm + 20 cm) (Figure 16). However, for the depth profile "litter or 1 + 20 cm", the worst case situation is "litter or 1 cm" because the concentration over 20 cm of course is lower than the concentration over 1 cm. Therefore, the depth profiles "litter or 1" and "litter or 1 + 20" could be merged to a worst case scenario of "litter or 1 cm", i.e., only two different worst case profiles.

A comparable rationale was used for defining depth profiles for enchytraeids (Table 11), reducing the 7 depth profile combinations to 3 worst case depth profiles: "litter or 1 cm", "2.5 cm", and "5 cm"(see also Figures 23 and 29).

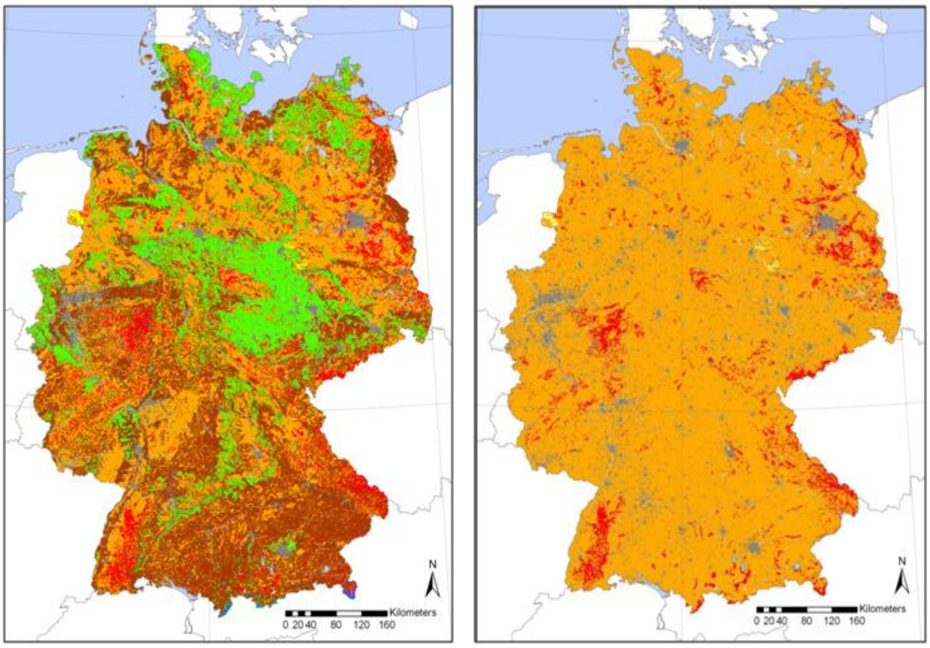


Figure 16: Example of the distribution maps of earthworm dominance classes (left hand) and the profile map (right hand) for Germany.

[19] Litter or 1 cm depth indicates that organisms are exposed via the litter layer, or if this is absent, via the upper layer of the soil (0 to 1 cm depth).

Table 10: Depth profile for earthworms (based on depths to model for each life form group). Soil depths considered are **1 cm** (0 to 1 cm) and **20 cm** (0 to 20 cm). In the case where a litter layer is present, the litter layer should be modelled instead of the 1 cm depth profile layer. Note that for anecic earthworms exposure also occurs via the burrows where they live, however this type of exposure is not yet considered in the present profiles. While epigeic earthworms are usually exposed via the litter layer, they can also occur on the soil surface (0 to 1 cm).

Dominance classes	**Distribution maps**	**Profile map (soil depths to model in cm depth)**	**Colour code map (Figure 28)**	**Worst case profiles (in cm depth)**
1	Anecic	Litter or 1	Red	Litter or 1
1 + 3	Anecic + epigeic	Litter or 1	Red	Litter or 1
3	Epigeic	Litter or 1	Red	Litter or 1
1 + 2	Anecic + endogeic	(Litter or 1) + 20	Orange	Litter or 1
2 + 3	Endogeic + epigeic	(Litter or 1) + 20	Orange	Litter or 1
1 + 2 + 3	All life forms	(Litter or 1) + 20	Orange	Litter or 1
2	Endogeic	20	Yellow	20

Table 11: Depth profile for enchytraeids (based on depths to model for each life form group). Soil depths considered are **1 cm** (0 to 1 cm), **2.5 cm** (0 to 2.5 cm), and **5 cm** (0 to 5 cm). In the case where a litter layer is present, the litter layer should be modelled instead of the 1 cm depth profile layer.

Dominance classes	**Distribution maps**	**Profile map (soil depths to model in cm depth)**	**Colour code map (Figure 29)**	**Worst case profiles (in cm depth)**
1	Litter dwellers	Litter or 1	Red	Litter or 1
1 + 2	Litter + intermediate dwellers	(Litter or 1) + 2.5	Red	Litter or 1
1 + 3	Litter + soil dwellers	(Litter or 1) +5	Red	Litter or 1
1 + 2 + 3	All life forms	(Litter or 1) + 2.5 + 5	Red	Litter or 1
2	Intermediate dwellers	2.5	Orange	2.5
2 + 3	Intermediate + soil dwellers	2.5 + 5	Orange	2.5
3	Soil dwellers	5	Yellow	5

4.6. Results

Note: the following results are based on presence/absence values of species, not on population sizes. This is important to note since differences in abiotic conditions often lead to shifts in dominance structure whereas the species or ecological group composition itself remains more or less unaltered. Whenever presented results appear counterintuitive this fact has to be taken into account.

4.6.1. Model outcome

In general the model fit shows reasonable results for the relative richness of all organism groups. The standardised residuals were checked for each influencing factor and showed no structure or asymmetry. Nevertheless, due to the high variation in the observed data the outcome revealed uncertainties at a local scale (e.g. a sample point or a field) since the results represent mean values on a larger regional scale. For more details refer to EFSA (in press).

4.6.1.1. Earthworms

In order to clarify the relationship between observations and the outcome of modelling, the number of sites classified in the seven life form dominance classes is presented in Table 12. As a general trend, it can be seen that the model "moves" sites from the outer borders of the triangle to its centre. For example, in the case of earthworms, 30 observed sites belong to the central life form category 123 (= all life forms occur at the respective site). According to the model, this number increases to 66 (Table 12). This increase reflects reality, since the importance of extreme observations for the general description of the occurrence of earthworm ecological groups decreases (for details, see Section 4.4.3).

According to the model developed for earthworms, considering the three countries together, the distribution of life forms was governed by the following explanatory variables: **Land use indicator, pH value, organic carbon content, precipitation, temperature, range of temperature**. In the following paragraphs, these factors are discussed in the light of data on the ecological preferences of earthworm species (and thus of ecological groups) known from literature (e.g. Binet et al., 1997; Briones et al., 1992; Curry, 1998; Edwards, 1998; Hendrix, 1998; Kladivko et al., 1997; Lavelle et al., 1997; Lee, 1985; Römbke et al., 2005). This discussion will help to support the results obtained with the maps.

Table 12: Concordance of observed and modelled life form dominance classes for adjusted relative richness of earthworms (203 sites)

No of Sites		Modelled Life Form Category							
		1	**2**	**3**	**12**	**13**	**23**	**123**	Sum
Observed Life Form Category	**1**	0	0	0	0	0	1	3	4
	2	0	7	1	0	0	25	12	45
	3	0	1	8	1	0	22	6	38
	12	0	2	0	1	0	5	17	25
	13	0	0	2	0	0	4	1	7
	23	0	1	3	1	0	26	11	42
	123	0	0	3	2	0	9	16	30
	Sum	0	11	17	5	0	92	66	191(a)

(a): For 12 sites were no fit possible

Land use

Land use is probably the most important factor determining the occurrence of earthworm species and/or ecological groups, despite the fact that vegetation per se is not correlated with the distribution pattern of earthworms: for example, at sites with the same pedological and climatic conditions but different plant cover (e.g. a deciduous forest and a grassland), the composition of the earthworm community is quite similar. However, differences in abundance or biomass do occur, which can be partly explained by differences in food quality or microclimatic conditions (e.g. soil temperature fluctuates more strongly in grassland sites than in a forest). Thus, the differences observed between the three land-use types are mainly due to human activities.

In grassland, the use of mineral fertilizers as well as soil compaction caused by heavy machinery or by cattle is probably causing stress for earthworms. In crop sites, mechanical factors strongly influence the earthworm community. Besides compaction, most detrimental is the direct damage to worms by tillage, the mechanical destruction of burrows by ploughing and the destruction of insulating litter layers. Irrigation usually positively influences earthworm populations, but this factor is certainly more important in arid and Mediterranean regions. An equally positive influence is caused by the input of organic material (such as farmyard manure, green manure or mulch), which not only increases the amount of food available, but also improves the soil properties in general (e.g. water regime, porosity etc.). However, some liquid organic manures can also be toxic. Certainly many plant protection products have detrimental effects on earthworms (e.g. some insecticides and fungicides), but the scale of the problem seems to decrease due to bans on very toxic and/or persistent compounds. In general, no-tillage cultivation, integrated and organic farming systems strongly favour earthworm populations, particularly large anecic species like *Lumbricus terrestris* (an ecosystem engineer).

pH value

In non-agricultural sites, the pH value is the most important factor governing the species distribution and the population size of earthworms. Therefore, species associations can be identified depending on the pH preference. However, when looking at agricultural sites, this statement is not completely true due to the fact that the soil at these sites has been anthropogenically changed. In Central Europe, and probably in most of the temperate regions, the pH of crop sites is between 5 and 6.5. Since most lumbricid species are ubiquitous or neutrophilic (i.e. occurring in a range of 4.5 – 7.0), the pH of crop sites is not usually a limiting factor for the composition of the earthworm community. However, even within this comparably narrow range, there is a pronounced increase of earthworm abundance with increasing pH. The situation in grasslands and orchards can be different, because many of these sites naturally have more acidic soils. Well-known examples are English grasslands at former moorland sites where the soil still has a quite low pH. Therefore, the occurrence of acid-tolerant species can be used to identify local factors responsible for a non-expected species composition or population pattern (e.g. a low number and biomass or a “strange” dominance spectrum).

Organic matter content

The distribution of earthworms is strongly influenced by the amount and quality of soil organic matter, since for (mainly, but not only) endogeic species this is the main food source. For example, in an English grassland their total number increased from 10 – 40 ind/m^2 at 1% carbon to 120 ind/m^2 at 2.5 % carbon content. Such a positive relationship was found for both endogeic and anecic species, while, as expected, epigeics do not react to changes in soil organic content. However, on a higher level of detail, organic matter content is difficult to use since additional factors become more important than the sheer content, e.g. how the organic matter originally reached the soil and, in particular, its quality, e.g. palatability. In addition, a high amount of organic matter in a certain soil can be either the result of a high input or be caused by slow degradation, for example in heavy metal contaminated soils; thus indicating very different living conditions for earthworms.

Precipitation

Little data is available directly showing the relationship between precipitation and earthworm distribution – more often the soil moisture is directly used. However, when analysing the influence of different factors on earthworms at various agricultural sites in Southern Australia, rainfall explained more of the variance in earthworm numbers than any other variable. Like in the case of temperature and soil moisture, temperature and rainfall are so closely inversely correlated that it is extremely difficult to identify which factor influences the earthworms most.

However, the overwhelming importance of soil moisture in determining earthworm distribution and activity is well known. This is true despite the fact that many earthworm species are very flexible concerning their moisture requirements (at least when their minimal demand is fulfilled; only extremes like poorly drained soils with low oxygen content are clearly avoided). On average, an actual soil moisture content higher than 20% for most of the year is sufficient to allow the development of a "good" earthworm community.

Temperature / Range of temperature

According to different authors, the optimum range of temperature for most soil-dwelling lumbricid earthworms in Central Europe is 10 - 18 °C. Below 10 °C many species do not reproduce anymore and below approximately 5 °C, most worms move to deeper soil layers, coil themselves up (summer or winter quiescence) or die (e.g. in non-covered agricultural soils in the US). Lumbricidae are known to survive in soils where surface layers are frozen, but usually die at temperatures not much below 0 °C. In a wheat field near Göttingen (Northern Germany) freezing of the uppermost bare soil (5 cm depth) was identified as the single most influential factor governing lumbricid population dynamics, in particular of endogeic species, but not of anecics. In the case of *Lumbricus terrestris*, only juveniles living close to the surface of a Swiss meadow died in winter when the temperature was below –5°C for a short period. Among the European species, *Dendrobaena octaedra* is the most cold-tolerant, being able to survive at –8 °C for three months and at –13.5 °C for two weeks in frozen soil. Cocoons of earthworm species living in cold regions have been shown to survive freezing temperatures.

Based on laboratory experiments, the upper lethal temperature is estimated as being 25 °C for *Eisenia fetida*, 26 °C for *Aporrectodea caliginosa* and 28 °C for *Lumbricus terrestris*; all values much lower than those measured at many sites where these species live. One has to keep in mind, though, that laboratory experiments are highly artificial. Under field conditions most earthworms will have the chance to avoid such extremes by migration to more favourable microsites (deeper in the soil or sheltered areas). Besides this, either populations adapted to low temperatures may have been used in these laboratory trials or the temperature effect was somehow mixed up with other factors like low soil moisture. For sure, temperature and moisture preferences are closely interrelated. In addition, populations of the same lumbricid species living in Europe or being introduced in, for instance, South Africa, are remarkably different concerning their temperature preference: *Aporrectodea rosea* in Germany: 12°C and in South Africa: 25 – 27°C.

Example of a factor NOT being significant: Texture

In general earthworms avoid soils with extreme textures, i.e., those with a high sand or clay content (probably also those which are not very favourable for agricultural purposes, at least cropped ones). In addition, they avoid pure sands due to the sharp-edged grains and often low content of organic matter as well as pure clays due to their oxygen deficiency. In other words, the available knowledge indicates that "normal" earthworm communities of agricultural lands thrive best in loamy soils. However, even in most other soils at least some earthworms will occur, meaning that in an investigation using presence/absence data this difference is smaller compared to studies based on abundance or biomass data.

4.6.1.2. Enchytraeids

As before, the relationship between the number of observed and modelled sites classified into the seven life form dominance classes of enchytraeids can be used to check the plausibility of the model (Table 13). However, in this organism group the distribution of life-form dominance is very robust, which might be due to the fact that the maximum numbers of species belonging to each of the three ecological groups are more evenly distributed among the three groups compared to earthworms (Table 7). Thus, the tendency of the model to “move” sites from the outer borders of the triangle to its centre is much weaker. In fact, for most of the classes the numbers are almost identical.

The following factors explained the distribution of life forms in enchytraeids: **pH value, organic carbon and clay content, temperature, range of temperature**. In the following paragraphs, these factors are discussed in the light of data on the ecological preferences of enchytraeid species (and thus of ecological groups) known from literature (e.g. Beylich and Graefe, 2009; Didden, 1993; Didden, 2002; Didden et al., 1997; Graefe and Beylich, 2003; Jänsch and Römbke, 2003; Langmaack et al. 1999). This discussion will help to support the results obtained with the maps.

Table 13: Concordance of observed and modelled life form dominance classes for adjusted relative richness of enchytraeids (159 sites)

No of Sites		**Modelled Life Form Category**							
		1	**2**	**3**	**12**	**13**	**23**	**123**	Sum
Observed Life Form Category	**1**	6	0	1	0	1	0	2	10
	2	0	0	1	0	0	0	0	1
	3	0	0	42	0	1	3	9	55
	12	0	0	0	2	0	3	8	13
	13	1	0	0	0	2	0	0	3
	23	1	0	4	0	1	0	2	8
	123	2	0	5	0	1	0	19	27
	Sum	10	0	53	2	6	6	40	117(a)

(a): For 42 sites no fits were possible

pH value

Like earthworms, enchytraeid species do have very different pH preferences. For example, Scandinavian coniferous forest sites or British moorlands, with pH-values clearly below 4, are totally dominated by just a couple of mainly litter-dwelling species such as *Cognettia sphagnetorum*. In any case, the species number is lower than 5 at acidic sites with thick litter layers. At the other end of the spectrum, soil-dwelling species belonging to the ecologically quite homogeneous genus *Fridericia* occur usually at sites with a pH > 4.2 (so far, only one litter-dweller has been identified in this genus: *Fridericia striata*). At higher pH values, at least until 6-7, no further change in species composition, and thus also the relationship between ecological groups, will occur.

Organic carbon content:

It is a well-known fact that enchytraeids regulate the activity of micro-organisms, both bacteria and fungi (especially in forest soils), thus being an important key group in the decomposition of organic matter and, indirectly, for maintaining the pool of organic matter in soil. However, there is hardly any information on the direct or indirect influence of enchytraeids on soil organic matter content. There is no general agreement whether soil microbes living on litter or the organic particles themselves are the most important food source for enchytraeids. As expected, soil-dwelling species have more mineral particles in their gut while litter-dwelling species prefer bacteria over fungi – but how rich organic matter content or quality influences the composition of enchytraeid ecological groups is still far from understood.

Clay content

The relationship between clay content and the distribution of enchytraeid ecological groups was only recently studied it detail. Data from German permanent soil monitoring sites (Bodendauerbeobachtungsflächen programme) suggest that a significant positive correlation between clay content and species number (mainly visible in an increase of *Fridericia* species) only exists in forests. For crops and grasslands no such correlation could be identified in a range between 2.5 and up to 40% clay. However, the number of grassland and crop sites investigated is still quite small.

Temperature / range of temperature

Except for some species regularly used in ecotoxicological laboratory tests, almost nothing is known about the preferred range of temperature even for the most abundant enchytraeid species. However, some information from the field is available. In temperate areas, low temperatures are typically the limiting factor during winter. On the other hand, the combination of low soil moisture and high temperatures can increase the desiccation stress for soil organisms, as the evaporation from the soil increases with temperature. One example is the highly dominant litter-dwelling species *Cognettia sphagnetorum*, which would certainly be restricted in its geographical distribution as well as in its functional activities above a threshold of 16°C. In general, the available information does not allow us to draw clear conclusions between climatic factors and the occurrence of enchytraeid ecological groups or even single species.

Example of a factor NOT being significant: Land use

An enchytraeid community in agricultural areas is expected to be very similar to a grassland community in its species composition, however reduced in abundance and number of species. This is due to the higher stress by soil cultivation. Both types of habitats are generally dominated by the genera *Fridericia* and *Enchytraeus*. At crop sites the dominance of intermediate species, mainly of the genus *Enchytraeus* is usually higher. At the same time, some of these species such as *Enchytraeus christenseni* appear as indicators of disturbance. At woodland sites the composition of species highly depends on the pH-value. In a very acidic spruce forest in Scheyern (Bavaria) totally different species were found compared to two adjacent agricultural and grassland sites. On the other hand, species in woodland with a higher pH value tend towards a grassland community. The number of species here may even exceed those found at grassland sites because typical woodland species like *Achaeta cf. affinis, Cognettia sphagnetorum, Stercutus niveus* or *Mesenchytraeus* sp. are also found (although only in small numbers). With these, a differentiation of grasslands and woodlands of higher pH-value is still possible. In summary, while there is clearly an influence of land use type on the abundance and species number of enchytraeids usually another factor mainly determines the composition of the enchytraeid community.

4.6.1.3. Collembola

The model effect of "pushing" sites to a more central position in the triangle (dominance of the three life form groups) was stronger for the Collembola (Table 14). This occurred mainly in sites where the observed values showed a dominance of single life form groups. This homogenisation of the relative richness among most sites indicates that the potential community composition is dominated by the three life form groups and may result in a weak separation within and between countries.

Table 14: Concordance of observed and modelled dominance classes for the raw relative richness of Collembola (250 sites)

No of Sites		**Modelled Life Form Category**							
		1	**2**	**3**	**12**	**13**	**23**	**123**	Sum
Observed Life Form Category	**1**	0	0	0	0	0	2	11	13
	2	0	0	0	0	0	9	10	19
	3	0	0	0	0	0	12	22	34
	12	0	0	0	0	0	2	6	8
	13	0	0	0	0	0	3	2	5
	23	0	0	0	0	0	47	42	89
	123	0	0	0	0	0	14	34	48
	Sum	0	0	0	0	0	89	127	216(a)

(a): For 34 sites no fit was possible

The model obtained for Collembola gave **Evapotranspiration** as the only significant factor explaining the distribution of life form classes from this group. However, the fit of the model was very poor, as can be seen by the missing concordance between the observed vs. the modelled categories of the dominant classes (Table 14). The most probable reason for this poor fitting is the spatial resolution of the environmental data used (1 km^2), which is too coarse to explain an ecologically meaningful spatial distribution pattern for this group with limited dispersal potential. Numerous other studies have demonstrated that, on a smaller scale that matches the sampling scheme, Collembola distribution can be very well explained by environmental factors like those used here (e.g. Filser et al., 2002). Another explanation for the poor fit could be the existence of other explanatory variables not considered in this exercise, such as bulk density. Also other variables related to management, in particular contamination, can have an important role in shaping Collembola communities (Filser et al., 1995).

4.6.1.4. Isopods

For isopods, the model presented a good plausibility check, with the observed and the modelled values of the adjusted relative richness indicating a similar profile of dominance classes (Table 15). This result was expected because in isopods only two life form groups were defined and one of them (the soil dwellers) presented a much lower number of species compared to the litter dwellers.

Table 15: Concordance of observed and modelled dominance classes for the adjusted relative richness of isopods (105 sites)

No of Sites		**Modelled Life Form Category**			
		1	**2**	**12**	Sum
Observed Life Form Category	**1**	1	2	1	4
	2	0	64	1	65
	12	0	7	1	8
	Sum	1	73	3	77(a)

(a): For 28 sites no fit was possible

The variables that were significant were **clay content**, **sand content, total precipitation and evapotranspiration**. The first two variables are related to the water holding capacity of the soil. So, together with total precipitation and evapotranspiration, all variables are closely related to soil water content. This is reasonable since the two life form groups of isopods were mainly defined on the basis of drought resistance of the species. Moreover, isopods are known to be highly susceptible to water loss and are thus restricted to moist habitats. Together with soil pH and calcium content, soil water content is one of the factors influencing most the distribution of isopods (Zimmer et al., 2000; Warburg et al., 1984).

4.6.2. Life-form composition in each country and mapping

4.6.2.1. Earthworms

Referring to Figure 17, the composition of earthworm life-forms differs considerably between the three selected countries. Most Finnish sites are not dominated by one ecological group but are characterized either by the occurrence of all three ecological groups or at least two, mostly endogeics and epigeics. Sites with only epigeics are quite common. Interestingly, one site is located almost at the far right side of the triangle, indicating a combination of endogeic and (mainly) anecic earthworms. In contrast, and not surprisingly due to the high number of sites investigated and the pedological and ecological heterogeneity of Germany, these sites are scattered across a wide area of the triangle. Most sites are characterized by the occurrence of all three ecological groups, with a strong tendency towards endogeic worms. Also some sites indicate the occurrence of just endogeic and anecics worms. Few sites are dominated either by endogeic or epigeic worms. Portuguese sites are characterized by epigeic and endogeic species, thus all sites are found at the left side of the triangle. The “relative richness” of the seven life-form dominance classes seems to be correct, in particular concerning the lack of sites dominated either by anecics alone or by a combination of anecics and endogeics. Thus, despite a considerable overlap in the occurrence of ecological groups of earthworms in the three countries the overall picture shows a characteristically different pattern.

Most of the area of [Southern] **Finland** (Figures 18 and 21) is dominated by epigeic species which is very much in line with the knowledge of Finnish soil properties and climate. Actually, the often very acid soils of this country, frequently in combination with coniferous forests (partly even moorland), are characteristic habitats for epigeic species of the genera *Dendrodrilus*, *Dendrobaena* and, partly, *Lumbricus*. However, in the most Southern region of Finland, there are also areas where anecic and endogeic species regularly occur, meaning that communities consisting of all three ecological groups as well as their combinations (endogeic–anecic as well as epigeic-anecic, but less often endogeic-epigeic) have been found regularly. This situation can be explained by the change in Finnish soil types but also by a change in land-use: in these Southern regions agriculture is regularly practiced while it does not make sense in Northern regions due to the acid, poor soils and long cold winters. In addition, lumbricid earthworms have had to resettle this land since the destruction of soils during the last ice-age about 10,000 years ago. In summary, the modelled distribution of Finnish earthworm ecological groups is reasonable.

The distribution of the earthworms in **Germany** is much more complicated (Figures 18 and 20). All combinations of earthworm ecological groups, except for communities consisting only of anecics, do occur, but in highly differing percentages. Regularly, the combination of endogeic and anecic species (e.g. *Aporrectodea caliginosa*, *Aporrectodea rosea* and *Lumbricus terrestris*) is found, mainly in grasslands. However, they can also be partly found at crop sites with higher clay contents, located in wide areas of Central Germany. Lastly, these species are found in some regions located close to the coasts. Also a combination of endogeics and epigeics, i.e. those *Aporrectodea* species already mentioned plus *Lumbricus rubellus* or *Lumbricus castaneus* is very common in almost all of Germany except in mountainous regions in the South. Here all life forms may occur together, consisting of the species listed plus one or two *Octolasion* species which, together with or alternatively to *Allolobophora chlorotica*, indicate wet conditions. However, sites with all ecological groups together are scattered all over Germany and may represent remnant deciduous forests on non-acid soils – a combination which is always lost in this intensively urbanized and agricultural country. As a proof of the robustness one may take the occurrence of sites dominated by epigeic species, which in fact are almost completely restricted to coniferous and beech wood forests with acidic soils, often in mountainous regions, but also at lowland sites in Northern Germany with poor acidic soils. The latter are more difficult to spot on the map since they rarely cover large areas. However, some pine plantations in former Eastern Germany are large enough to be seen even at this scale. Much rarer are sites with only endogeic species or the combination of anecic and epigeic species. For reasons still to be identified endogeic species are mainly located close to the Dutch border. So, while the distribution of earthworm ecological groups in Germany seems to make sense one should not forget that sampling efforts are very heterogeneously distributed in this country. While this is most obvious in Eastern Germany, the lack of data from the Northern German state of Schleswig-Holstein is an artefact: sampling was performed there, but the raw data was not available until very recently.

Finally, the distribution of earthworm ecological groups in **Portugal** resembles the Finnish situation with a clear differentiation between South and North. In the Southern part mainly endogeic species, partly together with epigeics, occur while the Northern part is dominated either by epigeics alone or in combination with endogeics (Figures 18 and 19). In fact, anecic species are very rare in this country, which may be caused by the often dry soils and high temperatures in the South and the usually dry and acidic soils in the North. An exception is the area around Lisbon which is characterized by a milder climate and richer soils. This, together with another land-use (agriculture) allows the occurrence of, for example, endogeic / anecic communities. It should be noted that Portugal, not being hit by glaciations, and despite its smaller size, has a higher biodiversity of earthworms than the other two countries – but this diversity is not apparent from the relatively low number of sampling sites, especially in the South.

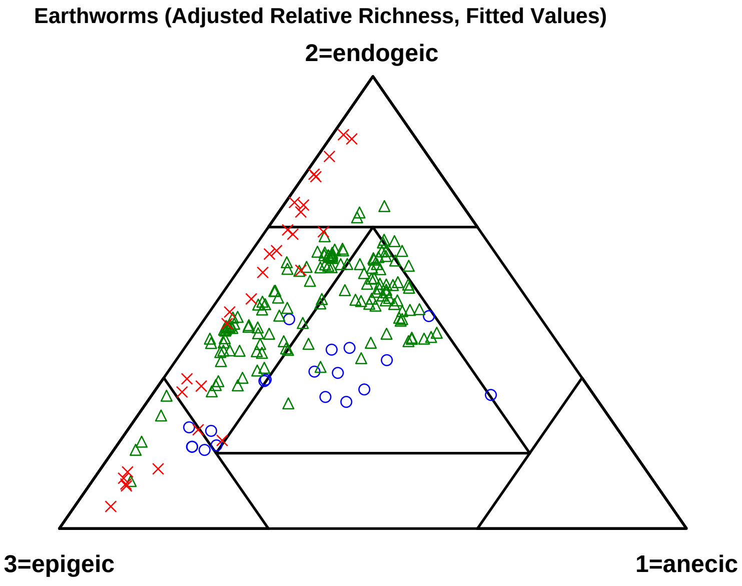


Figure 17: Life-form composition of earthworms (adjusted relative richness, fitted values) in the three countries Finland (blue circles), Germany (green triangles), Portugal (red crosses). For scale and categories see Figure 13.

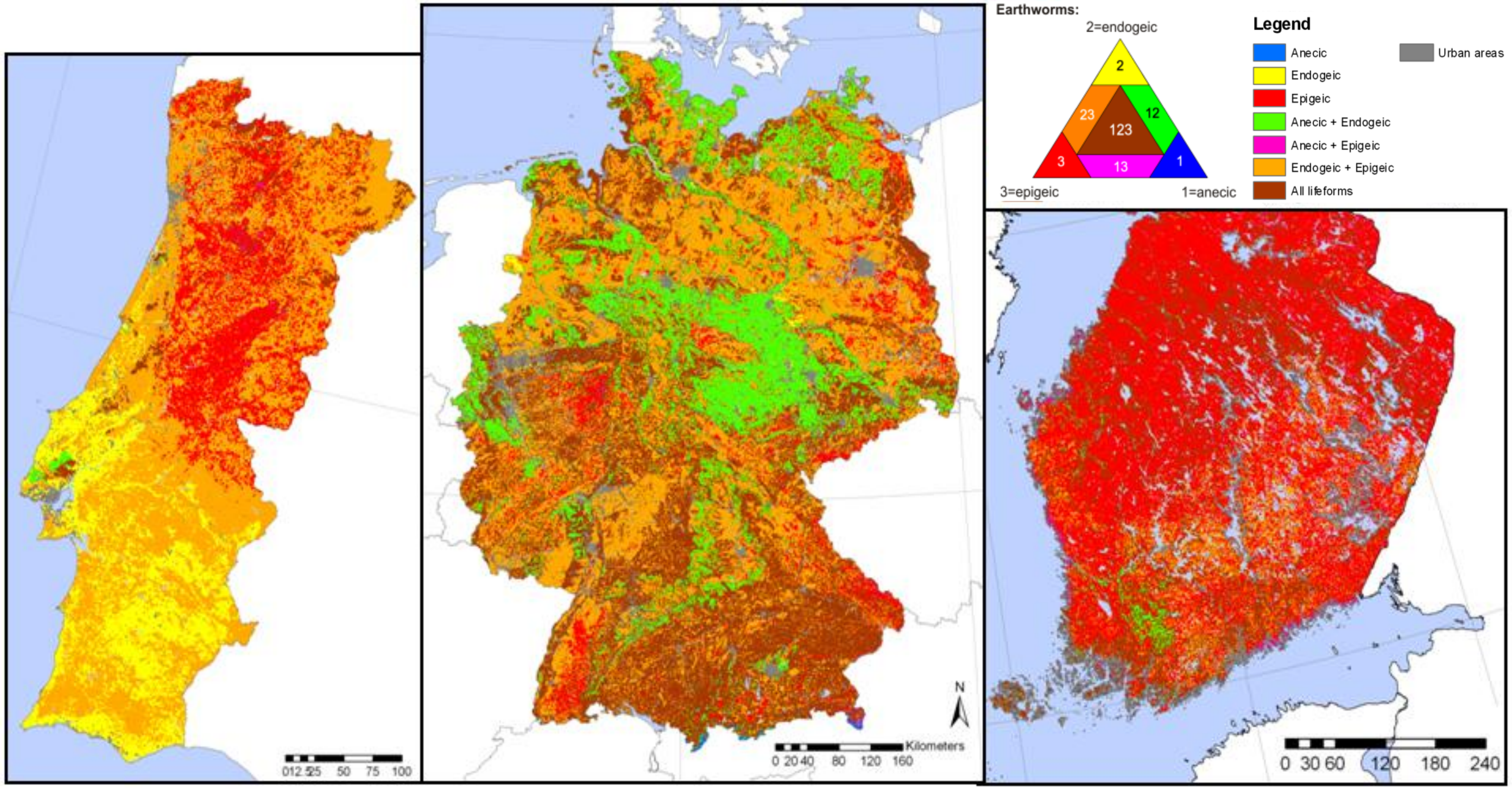


Figure 18: Maps showing the modelled distribution of earthworm life-form categories in the three countries Finland, Germany and Portugal. For Finland only the southern part of the country is shown because sampling data are only available there.

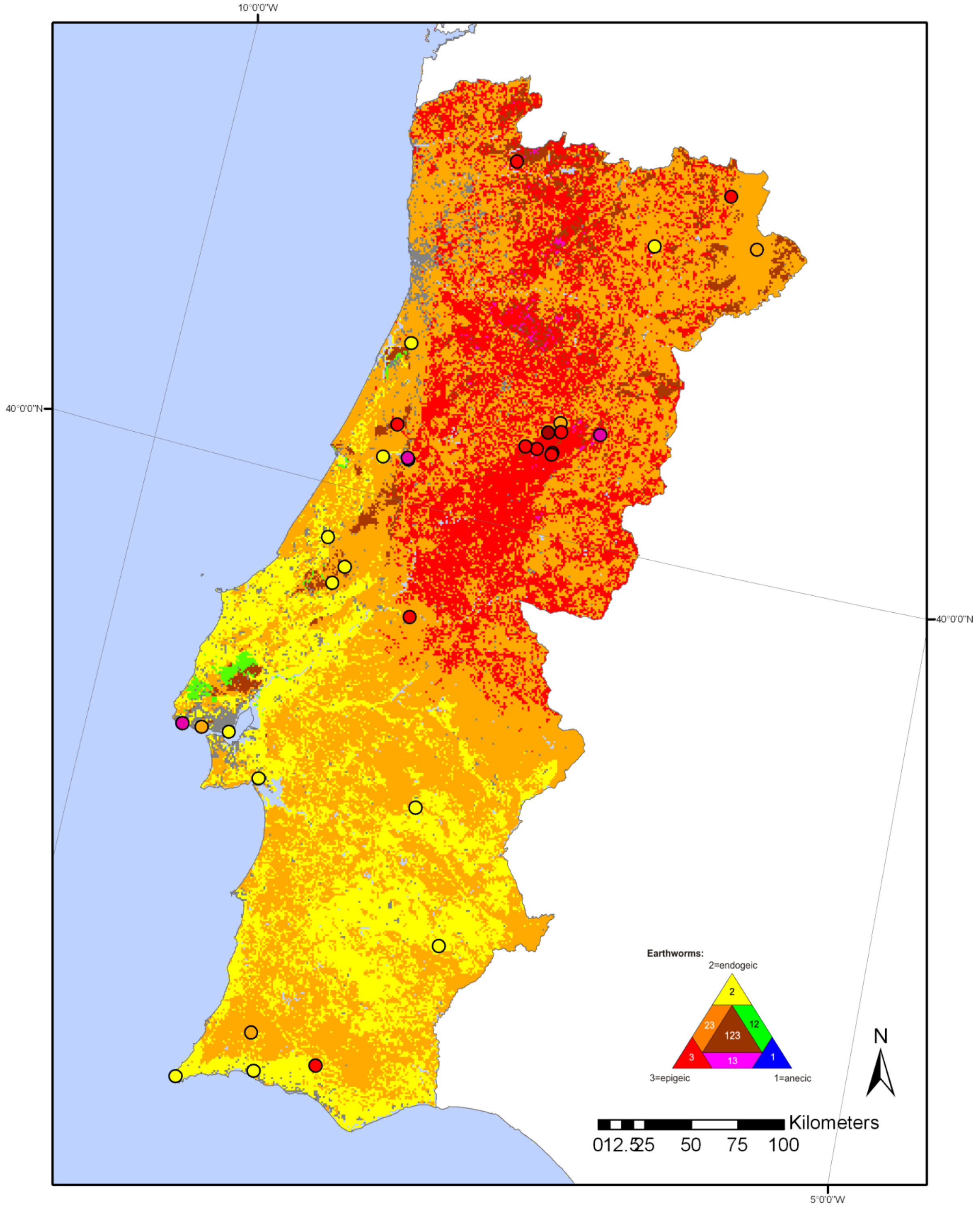


Figure 19: Map showing the modelled distribution of earthworm life-form categories in Portugal. Colours according to the classification scheme (grey = urban area). Single dots show the observations with their observed adjusted relative richness.

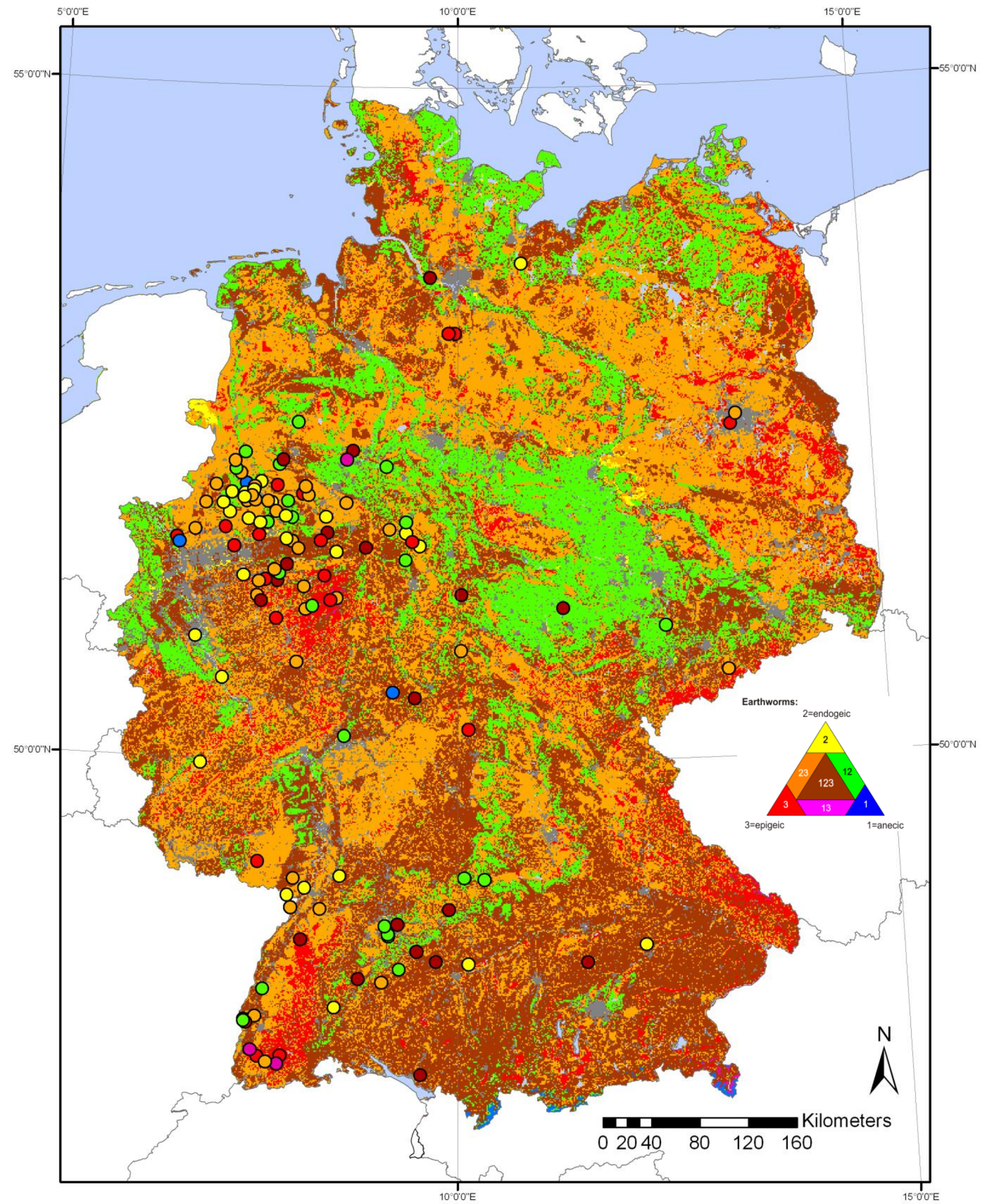


Figure 20: Map showing the modelled distribution of earthworm life-form categories in Germany. Colours according to the classification scheme (grey = urban area). Single dots show the observations with their observed adjusted relative richness.

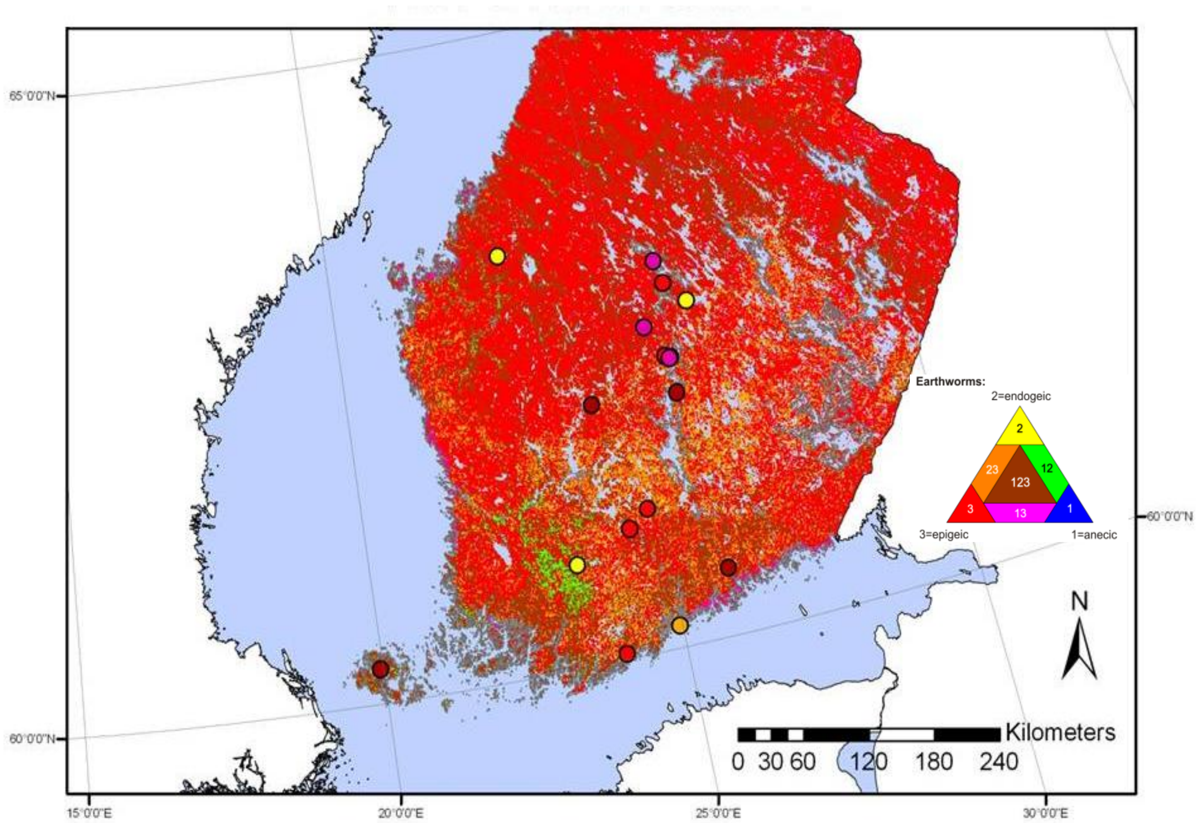


Figure 21: Map showing the modelled distribution of earthworm life-form categories in Finland. Only the southern part of the country is shown because sampling data are only available there. Colours according to the classification scheme (grey = urban area). Single dots show the observations with their observed adjusted relative richness.

4.6.2.2. Enchytraeids

The distribution of enchytraeid life-form dominance classes differs very strongly in the three selected countries (Figure 22). Almost all Finnish sites are highly dominated by litter-dwelling species. Litter- and soil dwelling species only occur together in few sites. Again, and probably for the same reasons, the German sites are scattered across a much wider area of the triangle. Roughly half of the investigated sites are dominated by soil-dwelling species or by a community with species from all three ecological groups. A few sites are characterised by species belonging to two different ecological groups, but not one site is dominated only by "intermediates". The distribution of Portuguese sites is not considered further since only one site has been studied so far. However, very recent developments indicate that there is a highly diverse enchytraeid fauna yet to be discovered in this country. Thus, the information provided does make ecological sense and indicates clear differences between Finnish and German sites in enchytraeid life-form patterns.

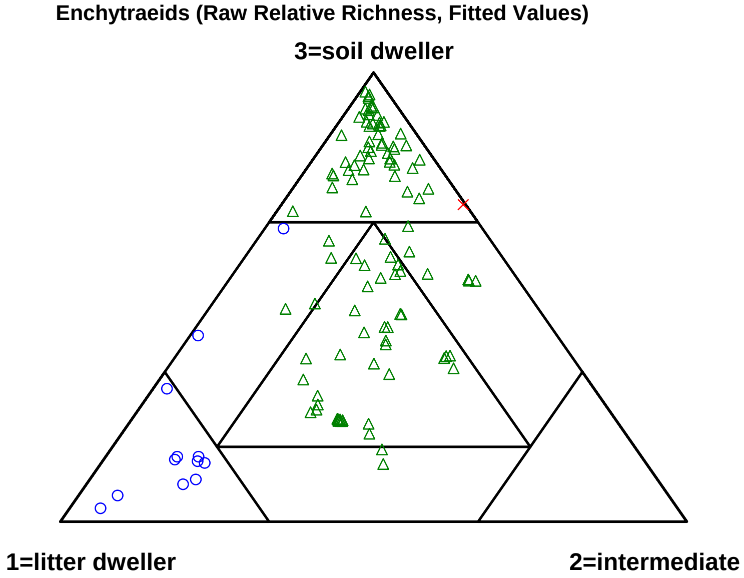


Figure 22: Life-form composition of enchytraeids (adjusted relative richness, fitted values) in Finland (blue circles), Germany (green triangles), Portugal (red crosses). For scale see Figure 13.

As already mentioned, the occurrence of enchytraeid ecological groups could only be modelled for the maps of Finland and Germany, since almost no data were available for Portugal. Despite a very limited number of sampling sites in the North of Finland, the whole country was used for modelling.

In the maps of **Finland** (Figures 23 and 25 at the end of this section), with a few exceptions in the South and some spots on the South-Eastern lake region, the country is dominated either by pure litter-dweller communities (especially in the North and East), pure soil-dwellers (mainly along the Western coast) or by combinations of both. Either soil or litter dwellers together with intermediates are quite isolated in the South, while the occurrence of all three ecological groups is restricted to the inner part of Finland in the southern third of the country, which is quite reasonable. However, it should not be forgotten that the model is based on presence/absence data. When abundance would have been included (which was not possible due to the lack of data), the dominance of litter-dwellers would be much higher: at many sites, the community is dominated by just one species (*Cognettia* *sphagnetorum*). Again, the whole western and northern parts of the country are clearly under-represented in terms of sampling effort, but even in the more complex areas of the south more research is needed in order to understand, for example, the occurrence of communities consisting of litter-dwellers and intermediates.

The modelled distribution of enchytraeid ecological groups in **Germany** looks at first sight quite uniform (Figures 23 and 24). Almost all of Western, Northern and Southern Germany is dominated by soil dwellers, partly together with intermediate dwellers. Also small areas dominated by epigeic and intermediate species occur regularly. This appears surprising but might be caused by the fact that at sites with partly quite different soil and site properties different species occur which however belong to the same ecological group: for example, in sandy and/or acid soils these are *Achaeta* species, while in richer, usually more neutral soils species of the genus *Fridericia* predominate. At the moment it is not clear why the situation is different in wide areas of Eastern Germany where intermediates and soil dwellers dominate the enchytraeid communities. Since a similar pattern is also observed in the Rhine valley as well as in the Wetterau (Hesse), it might be that higher temperatures are responsible for a shift in the dominance pattern of enchytraeid ecological groups. It seems surprising that almost no pure litter-dwelling communities are indicated, but this might be due to the fact already mentioned for the Finnish situation: here only presence/absence is considered– and, even at sites with a thick litter layer full of enchytraeids such as *Cognettia sphagnetorum*, there are at least a few, but sometimes very many, *Achaeta* or *Marionina* individuals living in deeper soil layers. Thus, the distribution pattern described here is in accordance with ecological and biogeographical knowledge on enchytraeids in Finland and Germany – but there are still some open questions, partly related to the lack of samplings in certain areas, partly caused by missing ecological information for several enchytraeid species.

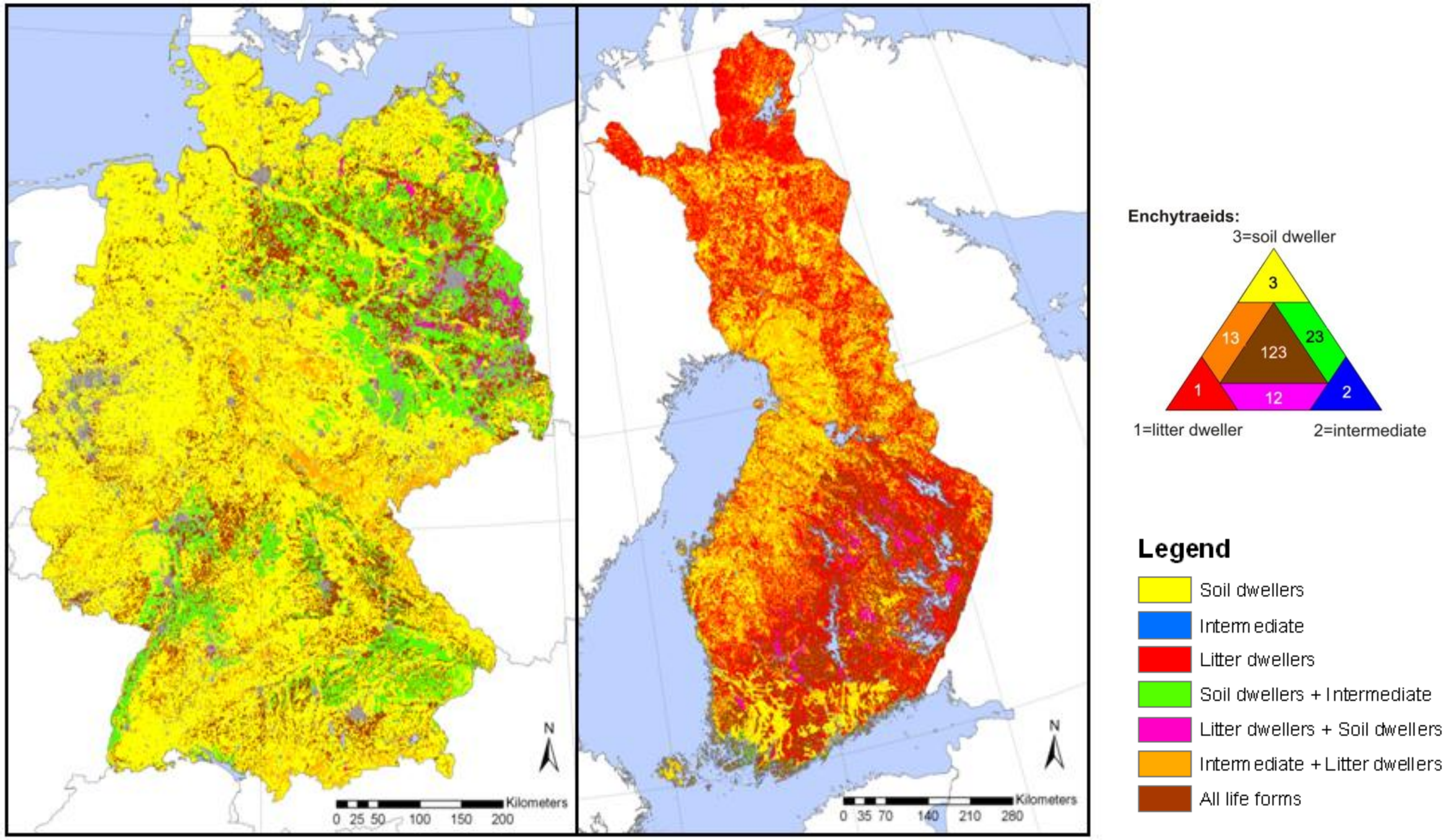


Figure 23: Maps showing the modelled distribution of enchytraeid life-form categories in Finland and Germany. Note that due to a lack of data it was not possible to draw up a map of the occurrence of enchytraeid life-forms in Portugal.

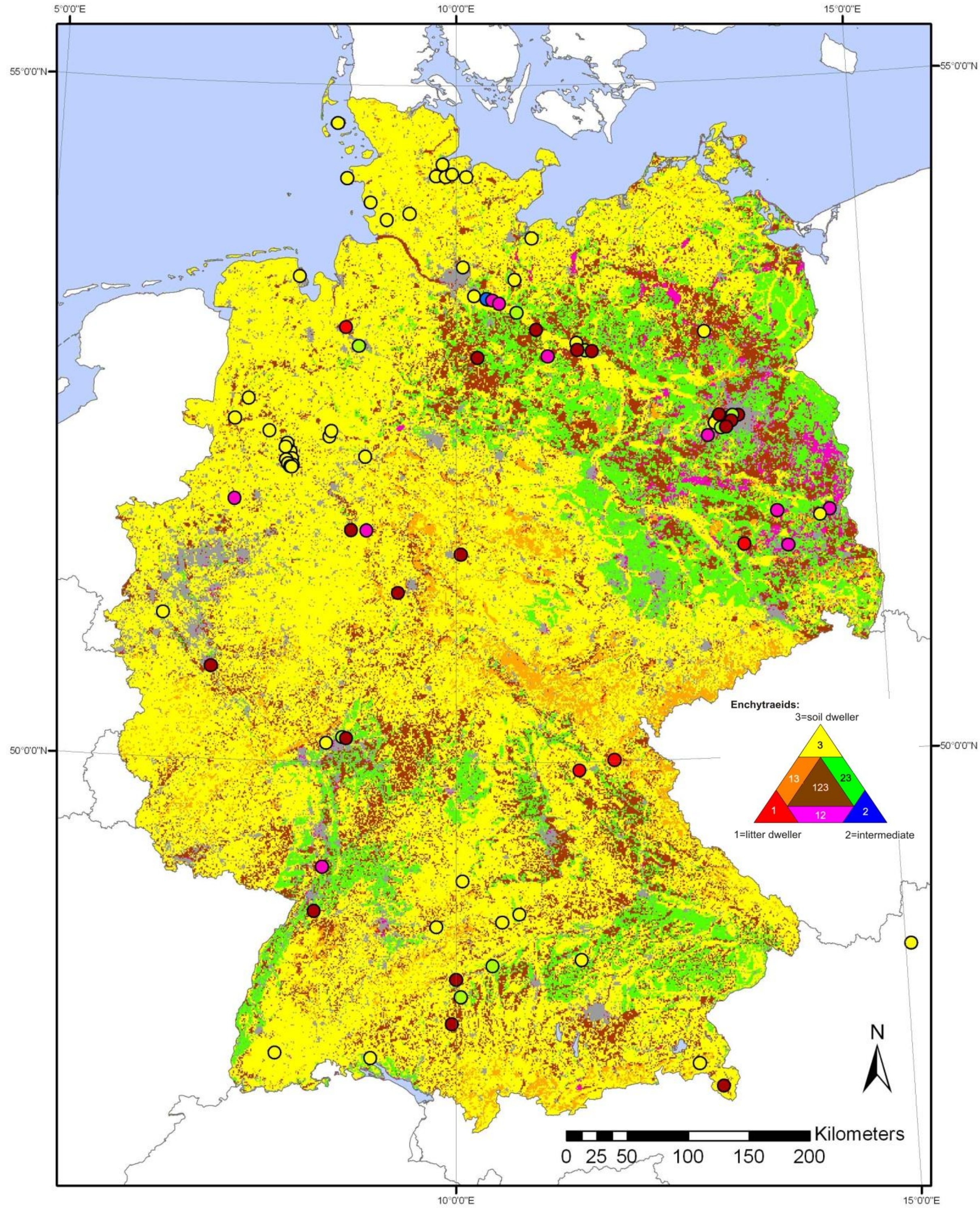


Figure 24: Map showing the modelled distribution of enchytraeid life-form categories in Germany. Colours according to the classification scheme (grey = urban area). Single dots show the observations with their observed adjusted relative richness.

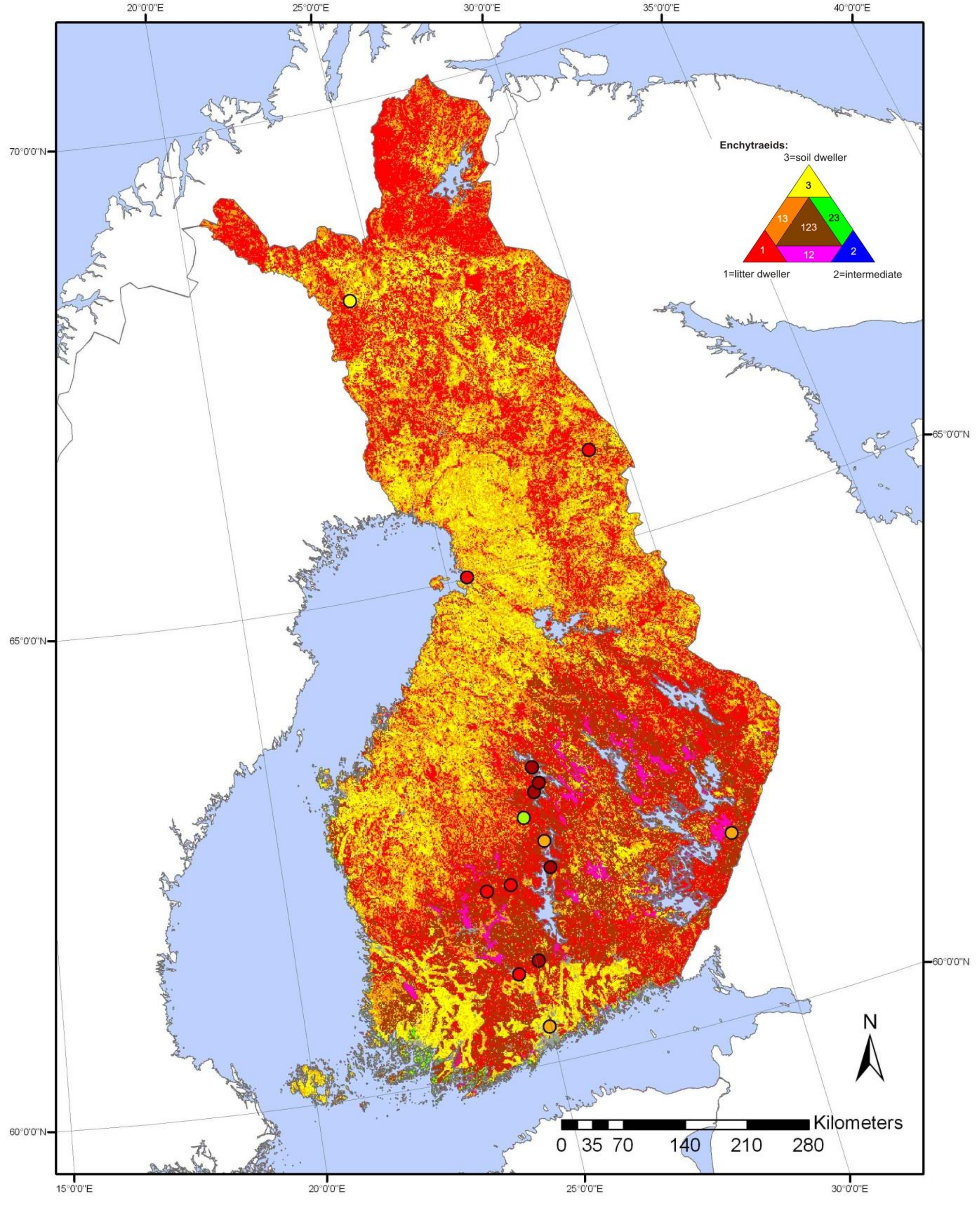


Figure 25: Map showing the modelled distribution of enchytraeid life-form categories in Finland. Colours according to the classification scheme (grey = urban area). Single dots show the observations with their observed adjusted relative richness.

4.6.2.3. Collembola

Collembola had the greatest diversity of all organism groups studied. The three life forms had a maximum number of 10, 16 and 24 (50 in total) species whereas in other groups the total species number amounted to roughly one third of this. Due to the high potential (and actual) diversity of Collembola in most sites representatives of all life forms were found. This even distribution resulted in the fact that the differentiation by the method used led to a rather poor discrimination both between and within countries. The model based on raw relative richness indicated a higher percentage of epigeic and hemiedaphic species in Portugal as compared to Germany (Figure 26), which seems somewhat counterintuitive with respect to climatic conditions. The Finnish data points are roughly located in-between, indicating more favourable conditions closer to the soil surface than in Germany, which could be expected based on expert knowledge.

This differentiation could be caused by the uneven distribution of data in the biogeographical database. When comparing data entries from forest and arable sites, their ratio amounts to roughly 20 in Portugal, compared to 5 and 1.4 in Finland and Germany, respectively. This means that forests are overemphasized in Portugal compared to the other two countries (especially Germany), which explains the differentiation between epigeic/hemiedaphic (more common in forest sites) and euedaphic species (more common in crop areas) between these two countries. Moreover, most Portuguese samples were taken from forests in the Northern Atlantic part of Portugal, presenting a well structured organic horizon, and soil samples were complemented with pitfall trap sampling, which contributed to emphasize the importance of epigeic species in Portugal in comparison to Germany. For a more balanced comparison of countries, adjusting the above mentioned ratio of forests and other land use types may reveal a more realistic differentiation. Moreover, as with enchytraeids, the actual species composition within one life form group also differs between countries.

Due to the reasons addressed in section 4.6.1.3, maps based on the modelled results did not show a convincing ecological meaning based on expert knowledge. Therefore, no maps are presented for Collembola.

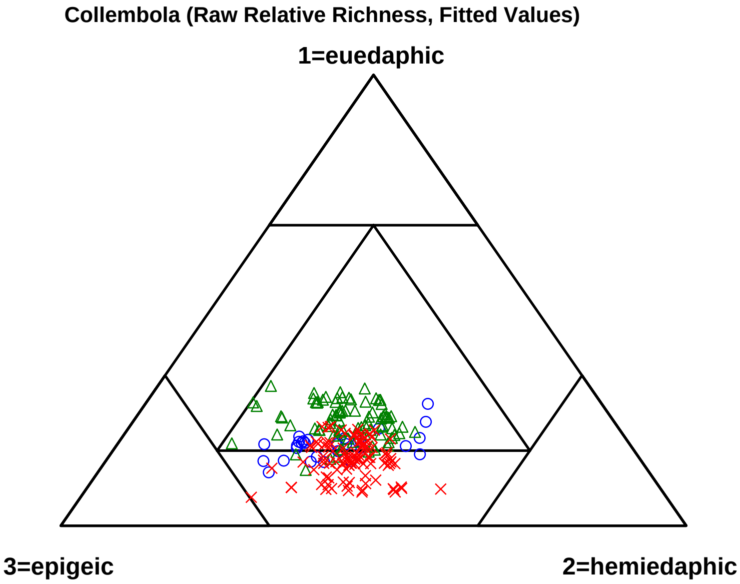


Figure 26: Life-form composition of Collembola (raw relative richness, fitted values) in the three countries. Finland (blue circles), Germany (green triangles), Portugal (red crosses). For scale and categories see Figure 13.

4.6.2.4. Isopods

Isopods were assigned to only two life form groups, resulting in the fact that all points in the triangle are situated on a line connecting species from the Trichoniscidae (soil dwellers) and all the other litter dwelling species (Figure 27). Trichoniscidae are more drought-sensitive than other groups of isopods and are found rather deeper in the soil (Matty Berg, VU Amsterdam, 2009). Despite this, the analysis gave no clear indication for patterns differing between or within countries. This is because only few species of this family were found and thus most model results revealed the litter dwellers as the dominant life form group. Therefore isopods were excluded from further analysis and are not shown as maps.

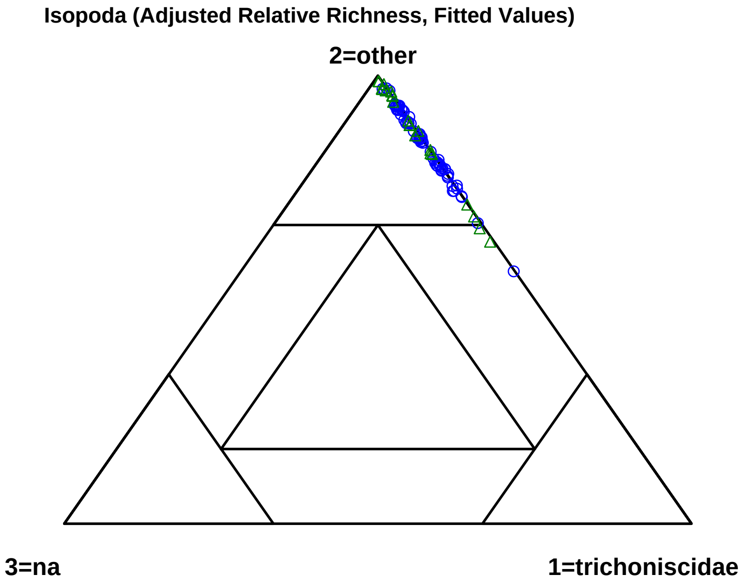


Figure 27: Life-form composition of isopods (adjusted relative richness, fitted values) in the two countries. Finland (blue circles), Germany (green triangles). For scale and categories see Figure 13. Soil information of data from Portugal are missing and therefore the data could not be included in the analysis.

4.6.3. Summary

The results described in the previous sections demonstrate that the approach used was able to show differences in the life form composition between the selected countries (a North-South transect) for both oligochaete groups (earthworms and enchytraeids) but not for the two arthropod groups (Collembola and isopods). For both earthworms and enchytraeids ecologically meaningful distribution maps could be produced, allowing the definition (as exemplified in the next section) of soil depth profiles where the ERCsoil could be modelled.

For the two arthropod groups the reasons why the approach that was used did not produce a meaningful response are different. For Collembola, the model gave a poor fit and produced, according to the existing ecological expertise, non reliable results. The solution may be better quality biogeographical data, including a balanced representation of land uses sampled and information on abundance of species, as well as environmental data at the local (site) scale.

For isopods the model approach gave a good fit. However this group was only separated into two life form classes, with one of them (the litter dwellers) containing most species and the other (soil dwellers) a very low number of species. Therefore the method failed to show a separation between sites, revealing the overall dominance of a single life form class. The definition of a third life form class would not result in a higher discriminatory power of this group since not only are most isopod species in fact litter dwellers, but also because any different traits used linked to habitat would lead to the same two division into two classes. Thus, the method worked, but for the reasons mentioned, isopods do not appear appropriate for defining soil ecoregions in the context of plant protection product risk assessment.

5. How to use the maps, definition of exposure profiles, consequences for environmental risk assessment

The maps issued based on biogeographical, climate and soil data (see Section 4.3), were "translated" into the relevant depth profiles for the respective soil organisms as summarized in this section. These profiles should be considered for the fate of plant protection products.

In Figures 28 and 29 the depth profiles for earthworms and enchytraeids are shown according to the depth profiles relevant for risk assessment.

For earthworms in the three studied countries (Finland, Germany and Portugal) (Figure 28), the areas coloured red and orange are both resulting in a worst case depth profile "litter or 1 cm" for risk assessment (Table 10). The worst case depth profile of "20 cm" (yellow surfaces on the maps) would only be applicable for 0.4 % of the surface area of Germany (Table 16). However, in Portugal, the "20 cm" worst case depth profiles (yellow surfaces in the map) are relevant in a significant part of Southern Portugal, amounting to 20.2 % of Portugal's surface area. For Finland, the worst case depth profile for the whole country is "litter or 1 cm".

For enchytraeids (Figure 29) the areas coloured red are resulting in a worst case depth profile "litter or 1 cm" for risk assessment (Table 11). Ares coloured orange result in a worst case depth profile of "2.5 cm" and those coloured yellow on a worst depth profile of "5 cm". In Germany the worst case depth profile for risk assessment in a larger surface of the country is the "5 cm" depth profile (yellow surfaces in the maps). In 15.4 % of the area, the worst case depth profile of "2.5 cm" would be applicable (orange surfaces on the maps) and in around 20% of the area (mainly in the former Eastern Germany), the worst case profile is "litter or 1 cm" (red surfaces on the map) is applied. However, in Finland, the "litter or 1 cm" worst case depth profiles (red surfaces in the map) are much more frequent than in Germany, reaching almost 80% of the surface, while the rest of the country would have a worst case depth profile of "5 cm" (20.9% - yellow surface of the map) or "2.5 cm (0.4% - orange surface on the map).

In line with the current risk assessment for plant protection products (EC, 2002) and the recommended approach of the PPR Panel (EFSA, 2010a, 2010d), the worst case soil depth profiles for short term risk assessment should be litter (if present), or 0 to 1 cm depth instead of the currently used 0 to 5 cm depth (Figures 28 an 29), while for refined risk assessments the geographical variation in depth profiles, crop and soil management information, as well as data about ecology of soil organisms (e.g. different dominance distribution of soil communities) could be considered. In addition, the use pattern of plant protection products, physico-chemical properties of the plant protection products, and the crop typologies could also be considered in the risk assessment (see also Figures 4 and 5 in EFSA, 2009; as well as EFSA, 2010a). The crop typology is particularly relevant considering the presence or absence of a litter layer in particular crops or agricultural cropping systems (e.g. low tillage) (EFSA, 2010b).When considering this information the risk assessment is more realistic and more focused. For example, if we focus on corn crops growing in areas where, based on expert knowledge or available local information, the dominance (in terms of abundance or biomass) of anecic earthworms is very low in comparison to endogeic species, one should switch from the 0-1 cm to the 0-20 cm depth, rendering a more realistic ERC soil value. Of course the opposite situation can also occur. Similar considerations can be made based on the use pattern of plant protection products and management practices for each crop type.

Table 16: Areas per earthworms and enchytraeid life form dominance classes and country (as km^2 and percentage of the country). The colour codes corresponds to the maps in the Figures 18 to 25 (life form classes) and to the maps in the Figures 28 to 29 (soil dept profiles), respectively.

EARTHWORMS								
Dominance classes		**Colour code maps**	**Germany**		**Finland**		**Portugal**	
			km^2	**%**	**km^2**	**%**	**km^2**	**%**
1	anecic	blue/ red	423	0.1	2484	0.8	0	0.0
3	epigeic	red/ red	22235	6.2	186771	57.5	17722	19.9
1 + 3	anecic + epigeic	Magenta/ red	189	0.1	36743	11.3	172	0.2
1 + 2	anecic + endogeic	green/ orange	59998	16.8	2532	0.8	349	0.4
2 + 3	endogeic + epigeic	orange/ orange	142799	40.0	21081	6.5	49457	55.6
1 + 2 + 3	all life forms	brown/ orange	129504	36.3	75406	23.2	3324	3.7
2	endogeic	yellow/ yellow	1483	0.4	0	0.0	17998	20.2
Total			**356632**	**100**	**325017**	**100**	**89023**	**100**

ENCHYTRAEIDS						
Dominance classes		**Colour code maps**	**Germany**		**Finland**	
			km^2	**%**	**km^2**	**%**
1	litter dwellers	Red/ red	311	0.1	112192	34.5
1 + 2	litter + intermediate dwellers	magenta/ red	7057	2.0	5821	1.8
1 + 3	litter + soil dwellers	orange/ red	17276	4.8	82032	25.2
1 + 2 + 3	all life forms	brown/ red	49126	13.8	55748	17.2
2	intermediate dwellers	blue/ orange	72	0.0	0	0.0
2 + 3	intermediate + soil dwellers	green/ orange	54883	15.4	1291	0.4
3	soil dwellers	yellow/ yellow	227907	63.9	67933	20.9
Total			**356632**	**100.0**	**325017**	**100.0**

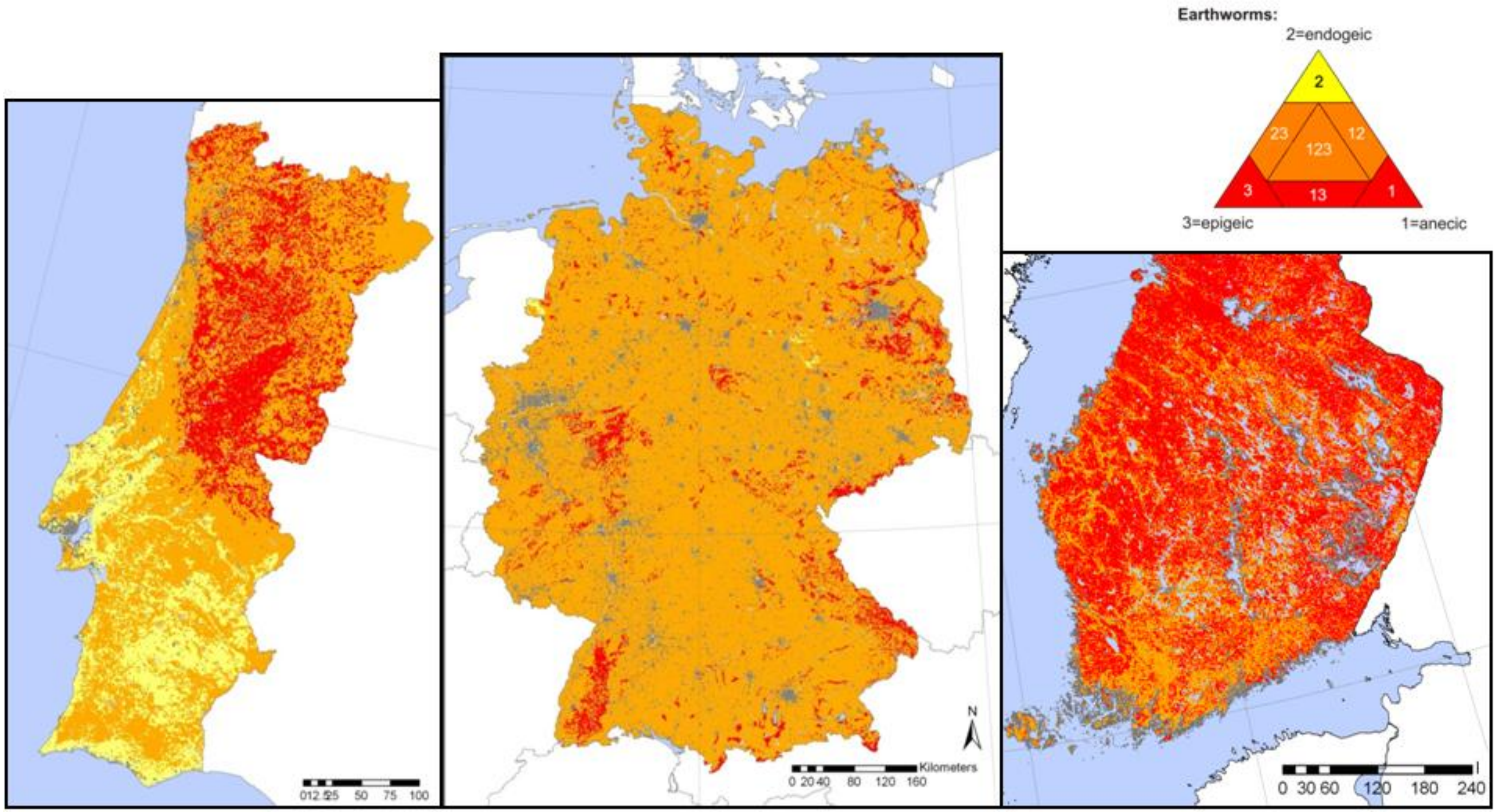


Figure 28: Earthworm profile maps, showing the depth profiles which should be modelled: **yellow** = 20 cm; **orange**= (litter or 1 cm) + 20 cm; **red** = litter or 1 cm. Please note the worst case profile for the orange regions would also be litter or 1 cm (as for the red regions).

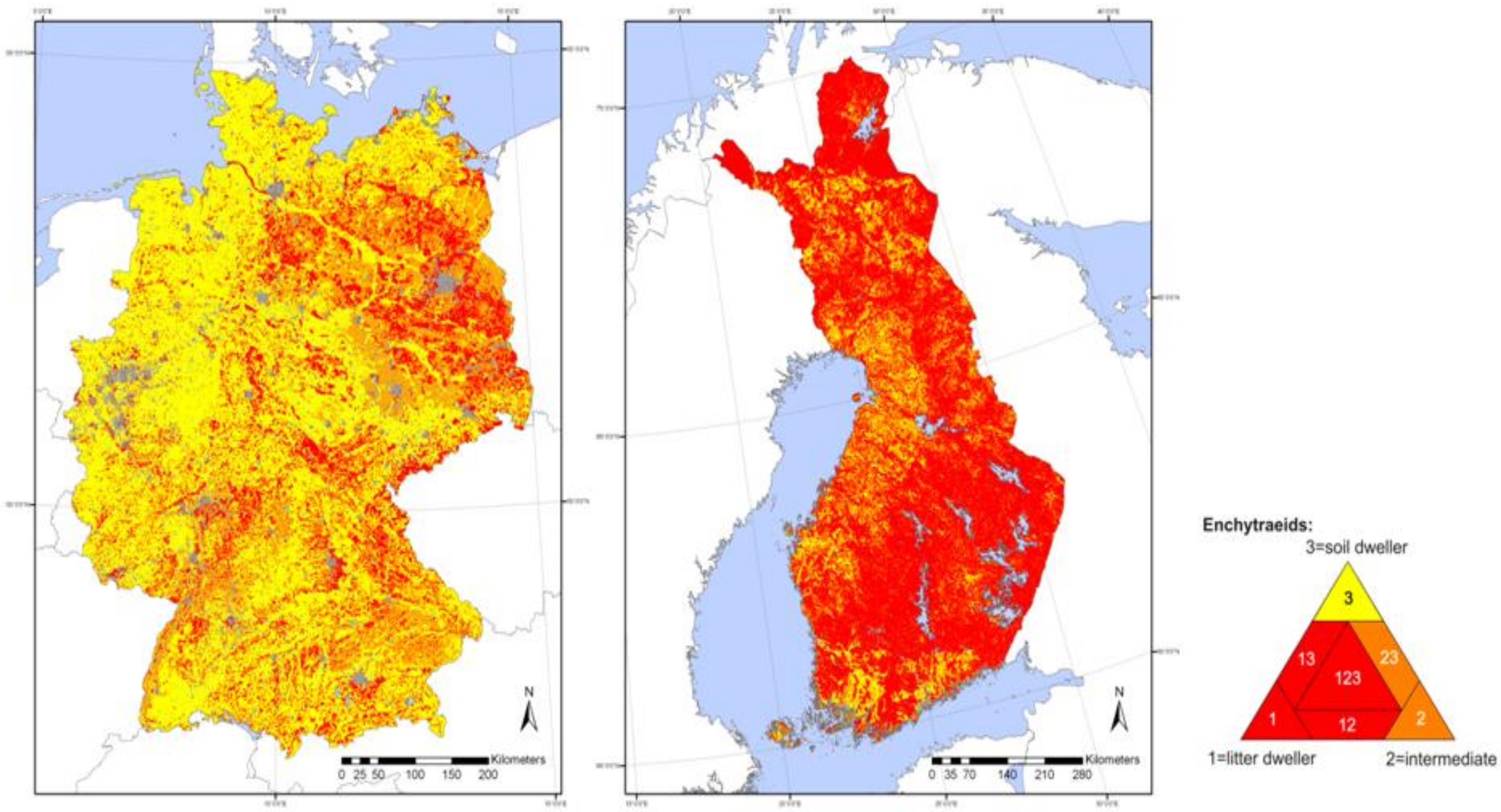


Figure 29: Enchytraeid profile maps, showing the depth profiles which should be modelled: **yellow** = 5 cm; **orange** = (litter or 1 cm) + 2.5 cm + 5 cm; **red** = litter or 1 cm. Please note the worst case profile for the orange regions would also be litter or 1 cm (as for the red regions).

CONCLUSIONS

The present approach (namely modelling of relative richness for certain soil organism groups) was performed for 3 countries (Finland, Germany, and Portugal) and four organism groups (earthworms, enchytraeids, collembolans, and isopods). It was possible to visualise life-form distributions for all organisms groups and to map ecoregions for earthworms and enchytraeids for most of the countries. Maps are not predictive on a local scale, but give a probability of the soil biota community to be found on a regional scale. The main results obtained are:

- Maps based on modelled information are in line with ecological and biogeographical information for the organism groups considered.
- Factors determining the distribution of the organisms could be identified, in particular for earthworms and enchytraeids.
- Differences could be observed between the three countries in community composition based on life form groups of earthworms and enchytraeids.
- The transformation of the information on the spatial distribution of the dominance classes into depth profiles for any ecotoxicologically relevant concentration (ERC) to be modelled was possible.

This procedure allows defining realistic "worst" case soil depth profiles for short term exposure. For realistic worst case, soil depth profiles of "litter or 1 cm" should be used for litter dwelling or epigeic organisms. For other life forms or for long term exposure, other depth profiles may represent the realistic worst case situation. For refined risk assessments the geographical variation in depth profiles, crop and soil management information, as well as data about ecology of soil organisms (e.g. different dominance distribution of soil communities) could be considered.

Provided that comparable information from other EU countries is available, the approach could be extended to the entire EU territory. However, several limitations and gaps in the availability of data constrained the analysis undertaken here:

- In comparison to grassland and forest sites, only a small number of studies in arable areas were available.
- The incomplete pedological description of the sampling sites led to the use of surrogate information on soil and climatic characteristics from additional databases, resulting in a mismatch of the resolution scales between biological and environmental data.
- The patchiness of the sampling sites led to situations where the information for different geographical areas was not homogeneous.
- Information on the abundances of species, which is more informative than presence / absence data used in the approach here, is limited.

A more complete data set would allow using more sophisticated models. In particular, a systematic data collection, the incorporation of biological knowledge on the influence of soil and climatic conditions on the occurrence of particular species (e.g. using non-linear functions), and the use of abundance data should significantly improve the outcome of this kind of analysis.

The original aim of this mandate was to define ecoregions by combining geographical information on different taxa. It proved possible to define ecoregion maps for earthworms and enchytraeids for most of the model countries.

The PPR Panel has demonstrated that the development of ecoregions is possible, provided that biogeographical data are available. The implementation of this approach would imply a refinement of the environmental risk assessment of plant protection products in soil which may need to be considered in the update of the Terrestrial Ecotoxicology Guidance Document (EFSA-Q-2009-00002) and, if appropriate, be discussed with risk managers. In particular, it seems that for most of the situations in Europe, the worst case soil depth profile for short term risk assessment would be litter (if present), or 0 to 1 cm depth instead of the currently used 0 to 5 cm depth.

It needs to be decided if the development of ecoregions should be expanded to cover the whole EU in the medium term, to improve the risk assessment in soil. This would imply producing and gathering further biogeographical data. Furthermore, the inclusion of additional taxa, in particular vulnerable species, could be considered.

RESEARCH NEEDS

An additional outcome was the identification of research needs addressing specific aspects of the approach described here:

- While the visualization by community triangles is recommended in general, it should be checked whether the outcome can be improved, e.g. by adjusting the size of the 7 classes.

- Modelling of exposure is so far limited to the occurrence of a plant protection product in depth layers of the mineral soil. However, certain organisms may not be exposed directly via the soil. This applies in particular to the litter layer (EFSA, 2010b).

- Since Collembola are known to be good indicators of soil conditions on a small scale, it should be studied if data sets collected on a smaller scale would improve the model fit, e.g. the data collected in several EU research projects such as the VULCAN project[20].

- Biological monitoring programs with standardised sampling methods focusing on agricultural areas should be performed on a regular basis, partly in order to fill data gaps in certain regions but also to collect better information concerning the "normal" ranges of species numbers, species composition, abundances and biomass (for general recommendations concerning biological soil monitoring programs see Römbke and Breure, 2005). This will improve the development of predictive models and reduce uncertainties.

- There is a general need to compile information on the biogeography and species composition of other soil organism groups not covered so far, e.g. mites, centipedes, diplopods, molluscs, nematodes and micro-organisms.

- When sampling soil organisms, it is recommended to also measure the most important soil and site parameters like pH, soil organic carbon, texture, cation exchange capacity and water-holding capacity as well as climatic factors (for details, see recommendations given in existing field test guidelines (e.g. ISO, 1999; OECD, 2006).

- More detailed geographical information on the extent of crop type and crop management practices such as tillage and/or irrigation is needed for risk assessment. Moreover, the way to integrate this kind of information with the biogeographical information needs to be developed.

[20] VULCAN - Vulnerability assessment of shrubland ecosystems in Europe under climatic changes. EU FP5 Contract EVK2-CT-2000-00094

DOCUMENTATION PROVIDED TO EFSA

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

EC (European Commission), 2002. Guidance Document on Terrestrial Ecotoxicology under Council Directive 91/414/EEC.

EFSA Mandate EFSA-Q-2009-00002. Consultation of the PPR Panel on the revision of the Guidance Document SANCO/10329/2002 (Terrestrial Ecotoxicology)

EFSA (European Food Safety Authority), 2009. Scientific Opinion of the Panel on Plant Protection Products and their Residues on a request from EFSA on the usefulness of total concentrations and pore water concentrations of pesticides in soil as metrics for the assessment of ecotoxicological effects. The EFSA Journal, 922, 1-90.

EFSA (European Food Safety Authority), 2009. Report on the PPR Workshop Improved Realism in Soil Risk Assessment (IRIS) – How will pesticide risk assessment in soil be tackled tomorrow? The EFSA Journal, 338, 1-32.

REFERENCES

Ad-hoc Boden AG, 2005. Bodenkundliche Kartieranleitung. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, 438 p.

ArcGis 9.3.1. ArcMap™ 9.3.1. ArcInfo © 1999-2009 ESRI Inc.

Beylich A and Graefe U, 2009. Investigations of annelids at soil monitoring sites in Northern Germany: reference ranges and time-series data. Soil Organisms, 81, 175–196.

Beck L, Dumpert K, Franke U, Mittmann H, Römbke J and Schönborn W, 1988. Vergleichende ökologische Untersuchungen in einem Buchenwald nach Einwirkung von Umweltchemikalien. Jülich Spezial, 439, 548-701.

Binet F, Hallaire V and Curmi P, 1997. Agricultural practices and the spatial distribution of earthworms in maize fields. Relationships between earthworm abundance, maize plants and soil compaction. Soil Biol. Biochem., 29, 577-583.

Bispo A, Cluzeau D, Creamer R, Dombos M, Graefe U, Krogh PH, Sousa JP, Peres G, Rutgers M, Winding A and Römbke J, 2009. Indicators for Monitoring Soil Biodiversity. Integrated Environmental Assessment Management, 5, 717-719.

Blume HP, Brümmer GW, Horn R, Kandeler E, Kögel-Knabner I, Kretzschmar R, Stahr K and Wilke BM, 2010. Scheffer /Schachtschabel: Lehrbuch der Bodenkunde. Spektrum Akademischer Verlag, Heidelberg, 570 pp.

Boesten JJTI, Köpp H, Adriaanse PI, Brock TCM and Forbes VE, 2007. Conceptual model for improving the link between exposure and effects in the aquatic risk assessment of pesticides. Ecotoxicology and Environmental Safety 66, 291-308.

Breure AM, Mulder M, Roembke J and Ruf A, 2005. Ecological classification and assessment concepts in soil protection. Ecotoxicology and Environmental Safety, 62, 211–229.

Briones MJI, Mascato R and Mato S, 1992. Relationships of earthworms with environmental factors studied by means of detrended canonical analysis. Acta Oecol., 13, 617-626.

Curry JP, 1998. Factors affecting earthworm abundance in soils. In: Edwards, C.A. (ed.): Earthworm ecology. CRC Press, Boca Raton, 37-64.

CSTEE (Scientific Committee on Toxicology, Ecotoxicology and the Environment), 2000. The available scientific approaches to assess the potential effects and risk of chemicals on terrestrial ecosystems. Opinion of the scientific committee on toxicity, ecotoxicity and the environment expressed at the 19th CSTEE plenary meeting. C2/JCD/csteeop/TER91100/D(0).

De Bello F, Lavorel S, Díaz S, Harrington R, Cornelissen JHC, Bardgett RD, Berg MP, Cipriotti P, Feld CK, Hering D, Silva PM, Potts SG, Sandin L, Sousa JP, Storkey J, Wardle DA and Harrison PA, 2010. Towards an Assessment of Multiple Ecosystem Processes and Services via Functional Traits. Biodiversity and Conservation, 19 (10), 2873-2893.

Didden WAM, 1993. Ecology of terrestrial Enchytraeidae. Pedobiologia 37, 2-29.

Didden WAM, 2002. A survey of enchytraeid communities in grassland and horticultural soil - The enchytraeid community as a monitoring tool. Newsletter on Enchytraeidae 7, 6-16.

Didden WAM, Fruend HL and Graefe U, 1997. Enchytraeids. In: Benckiser, G. (ed.): Fauna in soil ecosystems - Recycling processes, nutrient fluxes and agricultural production. M. Dekker, Inc. Ne.

Dunger W, 1978. Parameter der Bodenfauna in einer Catena von Rasen-Ökosystemen. Pedobiologia. 18, 310-340.

Dunger W, 1983. Tiere im Boden. A. Ziemsen Verlag, Wittenberg, 183 pp.

Dunger W and Dunger I, 1983. Zur Kongruenz von Phytozönosen und Collembolen-Synusien. Verh. SIEE X. Budapest, p. 32-34.

Dunger W, 1998. Die Bindung zwischen Bodenorganismen und Böden und die biologische Beurteilung von Böden. Bodenschutz 2, 62-68.

Edwards CA, 1998. Earthworm ecology. CRC Press, Boca Raton. 389 pp.

EC (European Commission), 2002. Guidance Document on Terrestrial Ecotoxicology under Council Directive 91/414/EEC (SANCO/10329/2002) rev.2 final, 17.10.2002, p.1 - 39.

EEA (European Environment Agency), 2000. Bossard M, Feranec J, Otahel, J. CORINE land cover technical guidance – Addendum 2000. Technical Report 40, Copenhagen .

EPPO (European and Mediterranean Plant Protection Organisation), 2003. EPPO Standards: Environmental risk assessment scheme for plant protection products. P/P 3 (revised), Soil. Bulletin OEPP/EPPO, Bulletin 33, 151-162.

EFSA (European Food Safety Authority), 2009. Scientific Opinion of the Panel on Plant Protection Products and their Residues (PPR Panel) on a request from EFSA on the usefulness of total concentrations and pore water concentrations of pesticides in soil as metrics for the assessment of ecotoxicological effects. The EFSA Journal, 922, 1-90.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2010a. Scientific Opinion on outline proposals for assessment of exposure of organisms to substances in soil. EFSA Journal 8(1): 1442. 38 pp.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2010b. Scientific opinion on the importance of the soil litter layer in agricultural areas. EFSA Journal, 8(6):1625. 21 pp.

EFSA (European Food Safety Authority), 2010c. Report on the PPR stakeholder workshop Protection goals for environmental risk assessment of pesticides: what and where to protect? EFSA Journal, 8 (7): 1672. 46 pp.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2010d. Scientific Opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the Guidance Documents on Aquatic and Terrestrial Ecotoxicology (SANCO/3268/2001 and SANCO/10329/2002). EFSA Journal, 8(10):1821. 55 pp.

EFSA, in press. Scientific report on the data analysis for the development of soil ecoregions using distribution data on invertebrates - feasibility study. EFSA Journal, in press.

Fierer N and Jackson RB, 2006. The diversity and biogeography of soil bacterial communities. Proc. Natl. Acad. Sci. USA 103:626–631.

Lauber CL, Strickland MS, Bradford MA and Fierer N, 2008. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. Soil Biol. Biochem. 40:2407–2415.

Filser J, Fromm H, Nagel R and Winter K, 1995. Effects of previous intensive agricultural management on microorganisms and the biodiversity of soil fauna. Plant and Soil, 170, 123-129.

Filser J, Mebes KH, Winter K, Lang A and Kampichler C, 2002. Long-term dynamics and interrelationships of soil Collembola and microorganisms in an arable landscape following land use change. Geoderma, 105, 201-221.

FOCUS (Forum for the Coordination of pesticide fate models and their use), 2000. FOCUS groundwater scenarios in the EU review of active substances. Report of the FOCUS Groundwater Scenarios Workgroup, EC Document Reference: SANCO/321/2000 rev.2, 202 pp.

Gardi C, Montanarella L, Arrouays D, Bispo A, Lemanceau P, Jolivet C, Mulder C, Ranjard L, Römbke J, Rutger SM and Menta C, 2009. Soil biodiversity monitoring in Europe: ongoing activities and challenges. European Journal of Soil Biology, 60, 807-819.

Graefe U and Beylich A, 2003. Critical values of soil acidification for annelid species and the decomposer community. Proc. 5th Int. Coll. Enchytraeidae. Newsletter on Enchytraeidae 8, 51-55.

Hendrix PF, 1998. Earthworms in agroecosystems: A summary of current research. In: Edwards, CA (ed.): Earthworm ecology. CRC Press, Boca Raton: 259-269.

ISO (International Organization for Standardization), 1999. Soil quality – Effects of pollutants on earthworms. Part 3: Guidance on the determination of effects in field situations. ISO 11268-3, Geneva, Switzerland.

Jänsch S and Römbke J, 2003. Ecological characterisation of selected enchytraeid species (Enchytraeidae, Oligochaeta) - A literature research. Proc. 5th Int. Coll. Enchytraeidae. Newsletter on Enchytraeidae 8, 57-68.

Kladivko EJ, Akhouri NM and Weesies G, 1997. Earthworm populations and species distributions under no-till and conventional tillage in Indiana and Illinois. Soil Biol. Biochem., 29: 613-615.

Langmaack M, Wiermann C and Schrader S, 1999. ‘Interrelation between soil physical properties and Enchytraeidae abundances following a single soil compaction in arable land’. Journal of Plant Nutrition and Soil Science, vol 162, pp517-525

Lauber CL, Hamady M, Knight R and Fierer N, 2009. Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil. Bacterial Community Structure at the Continental Scale. Applied and envirnomental microbiology, p. 5111–5120.

Lavelle P, Bignell D and Lepage M, 1997. Soil function in a changing world: the role of invertebrate ecosystem engineers. Eur. J. Soil Biol., 33: 159-193.

Lavelle P and Spain AV, 2001. Soil Ecology. Amsterdam: Kluwer Scientific Publications. 654p

Lee KE, 1985. Earthworms - Their ecology and relationships with soils and land use. Academic Press. 411 pp.

OECD (Organisation for Economic Co-Operation and Development), 2006. Guidance document on the breakdown of organic matter in litter bags. OECD Series on testing and assessment, 56. Environment Directorate, Paris, France.

Ranjard L, Dequiedt S, Jolivet C, Saby NPA, Thioulouse J Harmand J, Loise P, Rapaport A, Fall S, Simonet P, Joffre R, Chemidlin-Prévost Bouré N, Maron PA, Mougel C, Martin MP, Toutain B, Arrouays D and Lemanceau P, 2010. Biogeography of soil microbial communities: a review and a description of the ongoing french national initiative. Agron. Sustain. Dev. 30: 359–36.

Römbke J, Jänsch S and Didden W, 2005. The use of earthworms in ecological soil classification and assessment concepts. Ecotoxicology and Environmental Safety, 62, 249-265.

Römbke J and Breure AM, 2005. Status and outlook of ecological soil classification and assessment concepts. Ecotoxicology and Environmental Safety, 62, 300-308.

SAS Institute Inc. (2004): SAS/STAT 9.1 User’s Guide. Cary, NC: SAS Institute.

Warburg MR, Linsenmair KE and Bercovitz K, 1984. The effect of climate on the distribution and abundance of isopods. Symposia of the Zoological Society of London, 53, 339–367.

Zimmer M, Broll G and Topp W, 2000. Correspondence analytical evaluation of factors that influence soil macro-arthropod distribution in abandoned grassland. Pedobiologia, 44, 695-704.

APPENDICES

A. BIOGEOGRAPHICAL DATABASE STRUCTURE

The biogeographical database was compiled under a contract (CT/EFSA/PPR/2008/01[21]). The information in the database is divided into four major sections each one with several fields. The literature search was done on the Web of Knowledge (keywords "animal group" + "country"). Additionally, taxonomic based papers were searched browsing known scientific journals focusing on soil fauna that are not cited in the Web of Knowledge but were available to the contractor. Data from several projects reports and databases (e.g., the BIOASSESS project[22]) made available were also integrated in the database.

Section 1 – Site information (location, land-use)	
ID Entry	Entry in the database (one per each registry)
ID Site	Identification number of each site in the database
Country	Country name (in this case Finland, Germany, Portugal)
Region	Administrative region within the country.
Village/town	Name of the nearest village or town
Name/place	Name of the site
Coordinate (Long) Coordinates (Lat) Coordinates (format) Coordinates (datum)	Longitude Latitude Geographic system used (in this case UTM) Geodetic system used (in this case: WGS84) *NOTE: In the cases where no coordinates were mentioned in the literature searched, the coordinates were obtained using the approximate location of the sampling site (nearest village or town). This was done using the Google Earth search engine. In these cases the coordinates may fall within urban limits.*
Land-use	Land-use type (e.g., forest, pasture, crop area)
Dominant vegetation	Dominant vegetation at the site
Observations	Any relevant information can be placed in this field

[21] IMAR, 2009. Development of a biogeographical data base and of a draft European Ecoregion Map (EU 27), based on Lumbricidae, Collembola, Isopoda, and Enchytraeidae biogeographical and taxonomic literature.
[22] European Biodiversity Assessment Tools (BIOASSESS) EU FP5 project, Contract EVK4 -CT99-00280

Section 2 – Information on soil type and soil properties of that particular site	
ID Soil	Identification number of each soil type (usually at each site)
Class Class (typology used)	Soil class type and typology used
Texture Texture (typology used)	Soil texture and typology used
Sand (%) Silt (%) Clay (%)	Percentage of Sand, Silt and Clay
pH pH_SD pH_Min pH_Max	pH values (measure of variation and range if several pH values are reported for the same site)
Org. matter Org. matter_SD Org. matter_Min Org. matter_Max Org. matter_Unit	Soil organic matter content (measure of variation and range if several organic matter values are reported for the same site); Unit used (in most cases %)
Corg Corg_SD Corg_Min Corg_Max Corg_Unit	Soil organic carbon content (measure of variation and range if several soil organic carbon values are reported for the same site); Unit used (in most cases %)
Ntot Ntot_SD Ntot_Min Ntot_Max Ntot_Unit	Soil total nitrogen content (measure of variation and range if several Nitrogen values are reported for the same site); Unit used (in most cases %)
C/N C/N_SD C/N_Min C/N_Max	Soil C/N ratio (measure of variation and range if several C/N values are reported for the same site)
WHCmax WHC_Unit	Soil maximum water holding capacity and unit of expression
Humus type Reference Humus type	Humus type (if mentioned) typology used
Observations	Any relevant information can be placed in this field

Section 3 – Information on the species	
ID Sp	Identification number of each species in the database
Order, Family, Species Author, Year	Taxonomic information (including author and year of description)
Life-form typus	Information on life-form (dependent of the organism group). For **Collembola** 3 morphological traits were used to define the life form: ocelli, antenna and furca. Each one was coded between 1 and 5 as follows: Ocelli: 1= (0+0) ocelli; 2= (1+1)-(2+2) ocelli; 3=(3+3)-(4+4) ocelli; 4=(5+5)-(6+6) ocelli; 5=(7+7)-(8+8) ocelli Antenna: 1= <0.25 of body length; 2= 0.25-0.5 of body length; 3= 0.5-0.75 of body length; 4= 0.75-1 of body length; 5= >1 of body length Furca: 1= absent; 3= reduced/short; 5= fully developed These three traits were combined to create the following life-form typology (an higher score indicates a life–form adapted to upper soil layers and with a high dispersal capability): Life-form classes: class 1= score 1-3 (euedaphic; very low dispersal); class 2= score 4-6 (euedaphic-hemiedaphic; low dispersal); class 3= score 7-9 (hemiedaphic; medium dispersal); class 4= score 10-12 (hemiedaphic-epigeic; medium-fast dispersal); class 5= score 13-15 (epigeic-fast dispersal) For **earthworms** 3 life-form traits were considered: Anecic - species that live in permanent vertical burrows in mineral soil layers (up to 3 m deep) Endogeic - species that inhabit mineral soil, making horizontal non-permanent burrows, mainly in the uppermost 10 – 20 cm of soil Epigeic - species that live above the mineral soil surface, typically in the litter layers of forest soils (partly on tree bark), making no burrows For **enchytraeids** 3 life-form traits were considered: Soil dwellers - species that live mainly in soil (up to 5 cm depth) Intermediate dwellers - species that circulate between soil and litter layers Litter dwellers - species that live mainly in the litter layer
Depth	Soil depth at which the species was collected
Horizon	Horizon (litter layer or soil) at which the species was collected
Abundance Abundance_Min Abundance_Max Abundance_basis	Abundance of the species in the set of samples or sampling date. For Collembola this value can vary: total nº of individuals in the sample (the default measure); nº individuals/m2; nº individuals/trap (in case of the use of pitfall traps). For earthworms and enchytraeids this value is usually given in nº individuals/m2. For isopods this value (when available) is usually given in nº individuals/trap (in case of the use of pitfall traps)
Sampling method	Sampling method used to collect samples
Observations	Any relevant information can be placed in this field. For earthworms and enchytraeids dominance data is given in this field.

Section 4 – Information on the source of the data (database, publication, report)	
ID Ref	Identification number of each data source
First author	Name of the first author
Journal / Source	Name of the data source (usually a journal)
Year	Year of publication
Volume (issue)	Volume & issue (when applicable)
Pages	Page numbers (when applicable)
Observations	Any relevant information can be placed in this field

B. Soil and climate data used in the data analysis and mapping

All data are provided with a common projection (ETRS 89 LAEA) and use the same spatial frame with a resolution of 1000 m (Figure 30). The file format is ESRI ARC-Grid raster and data are stored in ASCII.

Common Properties

file format	: ESRI ARCRASTER ASCII
columns	: 3500
rows	: 4100
ref. system	: ETRS 89 LAEA
unit dist.	: 1.0000000
min. X	: 2500000.0
max. X	: 6000000.0
min. Y	: 1412000.0
max. Y	: 5512000.0
resolution	: 1000
unit	: m

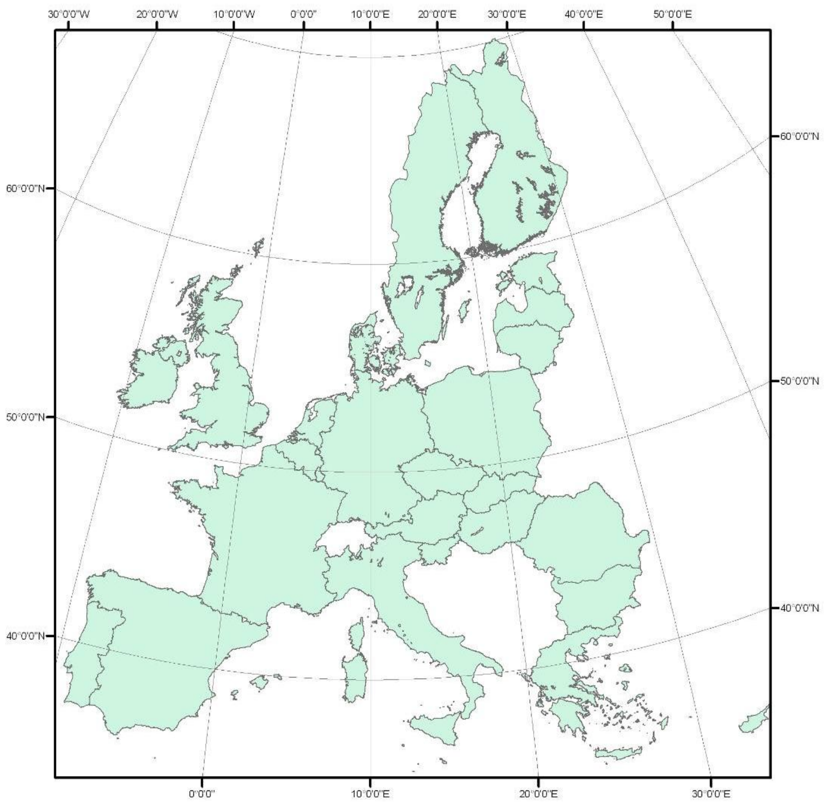


Figure 30: Spatial extent of the provided raster maps.

Organic Carbon in Topsoil

File name	: OC_EFSA
Layers	: 1
file type	: real
data type	: real
ref. units	: %
value units	: %
flag value	: 0
flag def'n	: background
Source	: JRC OC_TOP
Processing	: R. Hiederer, JRC
Reference	: Jones, R.J.A, R. Hiederer, E. Rusco, P.J. Loveland and L. Montanarella (2005). Estimating organic carbon in the soils of Europe for policy support. European Journal of Soil Science, October 2005, 56, p.655-671.

Topsoil pH

File name	: PH_TOP_EFSA
Layers	: 1
file type	: real
data type	: real
ref. units	: none
value units	: pH in water
flag value	: 0
flag def'n	: background
Source	: http://www.iiasa.ac.at/Research/LUC/luc07/External-World-soil-database/HTML/
Processing	: R. Hiederer, JRC
Reference	: FAO/IIASA/ISRIC/ISS-CAS/JRC, 2008. Harmonized World Soil Database (version 1.0). FAO, Rome, Italy and IIASA, Laxenburg, Austria. 37pp.

Soil Texture

File name	: Text_DOM_EFSA
Layers	: 1
file type	: integer
data type	: binary
ref. units	: none
value units	: soil texture class
flag value	: 0
flag def'n	: background
Source	: http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/index.htm
Processing	: R. Hiederer, JRC
Reference	: SOIL GEOGRAPHICAL DATABASE OF EURASIA AT SCALE 1:1,000,000, VERSION 4 beta, 25/09/2001

Legend:
1 Coarse (18% < clay and > 65% sand)
2 Medium (18% < clay < 35% and >= 15% sand, or 18% <clay and 15% < sand < 65%)
3 Medium fine (< 35% clay and < 15% sand)
4 Fine (35% < clay < 60%)
5 Very fine (clay > 60 %)
9 No mineral texture (Peat soils)

Corine Land Cover 2000

File name	: CLC_2000_EFSA
Layers	: 1
file type	: integer
data type	: binary
ref. units	: none
value units	: land cover class
flag value	: 0
flag def'n	: background
Source	: Corine land cover 2000 (CLC2000) 250 m - version 8/2007 http://www.eea.europa.eu/themes/landuse/clc-download
Processing	: R. Hiederer, JRC
Reference	: M.V. Nunes de Lima (2005) European Commission Joint Research Centre, Institute for Environment and Sustainability (IES), Land Management Unit, I-21020 Ispra (VA), Italy. EUR 21757 EN, ISBN 92-894-9862-5.

Legend:

GRID	CLC	LABEL3
1	111	Continuous urban fabric
2	112	Discontinuous urban fabric
3	121	Industrial or commercial units
4	122	Road and rail networks and associated land
5	123	Port areas
6	124	Airports
7	131	Mineral extraction sites
8	132	Dump sites
9	133	Construction sites
10	141	Green urban areas
11	142	Sport and leisure facilities
12	211	Non-irrigated arable land
13	212	Permanently irrigated land
14	213	Rice fields
15	221	Vineyards
16	222	Fruit trees and berry plantations
17	223	Olive groves
18	231	Pastures
19	241	Annual crops associated with permanent crops
20	242	Complex cultivation patterns
21	243	Land principally occupied by agriculture, with significant areas of natural vegetation
22	244	Agro-forestry areas
23	311	Broad-leaved forest
24	312	Coniferous forest
25	313	Mixed forest
26	321	Natural grasslands
27	322	Moors and heathland
28	323	Sclerophyllous vegetation
29	324	Transitional woodland-shrub
30	331	Beaches, dunes, sands
31	332	Bare rocks
32	333	Sparsely vegetated areas
33	334	Burnt areas
34	335	Glaciers and perpetual snow
35	411	Inland marshes
36	412	Peat bogs
37	421	Salt marshes
38	422	Salines
39	423	Intertidal flats

40	511	Water courses
41	512	Water bodies
42	521	Coastal lagoons
43	522	Estuaries
44	523	Sea and ocean
48	999	NODATA
49	990	UNCLASSIFIED LAND SURFACE
50	995	UNCLASSIFIED WATER BODIES

Mean Monthly Temperature 1960-1990

File name	: TMEANn_EFSA
Layers	: 12
file type	: real
data type	: real
ref. units	: deg C
value units	: average monthly mean temperature
flag value	: -9000
flag def'n	: background
Source	: http://www.worldclim.org/
Processing	: R. Hiederer, JRC
Reference	: Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978

Min Monthly Temperature 1960-1990

File name	: TMin_EFSA
Layers	: 12
file type	: real
data type	: real
ref. units	: deg C
value units	: average monthly mean temperature
flag value	: -9000
flag def'n	: background
Source	: http://www.worldclim.org/
Processing	: R. Hiederer, JRC
Reference	: Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978

Max Monthly Temperature 1960-1990

File name	: TMax_EFSA
Layers	: 12
file type	: real
data type	: real
ref. units	: deg C
value units	: average monthly mean temperature
flag value	: -9000
flag def'n	: background
Source	: http://www.worldclim.org/
Processing	: R. Hiederer, JRC
Reference	: Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978

Mean Annual Precipitation 1960-1990

File name	: PMEANn_EFSA
Layers	: 1
file type	: real
data type	: real
ref. units	: mm
value units	: Mean Annual precipitation
flag value	: -9000
flag def'n	: background
Source	: http://www.worldclim.org/
Processing	: R. Hiederer, C. Gardi, JRC
Reference	: Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978

Mean Annual Potential Evapotranspiration 1975-2009

File name	: ET0_EFSA
Layers	: 1
file type	: real
data type	: real
ref. units	: mm
value units	: Mean Annual Potential Evapotranspiration (Penmann)
flag value	: -9000
flag def'n	: background
Source	: http://mars.jrc.it/mars/
Processing	: C. Gardi, JRC
Reference	: Erik van der Goot, November 1997; Stefania Orlandi, December 2003. Technical description of interpolation and processing of meteorological data in CGMS. Available at:http://mars.jrc.it/mars/About-us/AGRI4CAST/Data-distribution/Data-Distribution-Grid-Weather-Doc

C. Procedures for the preparation of the Soil Organisms Geographical Database and Data Extraction

The original biogeographical database provided for the three test countries (Finland, Germany and Portugal) was organized in separate Excel spreadsheets (for the different groups of soil organisms), and the geographic coordinates were based on UTM (Universal Transverse of Mercator) coordinate system, based on Datum WGS 84 (World Geodetic System 1984).

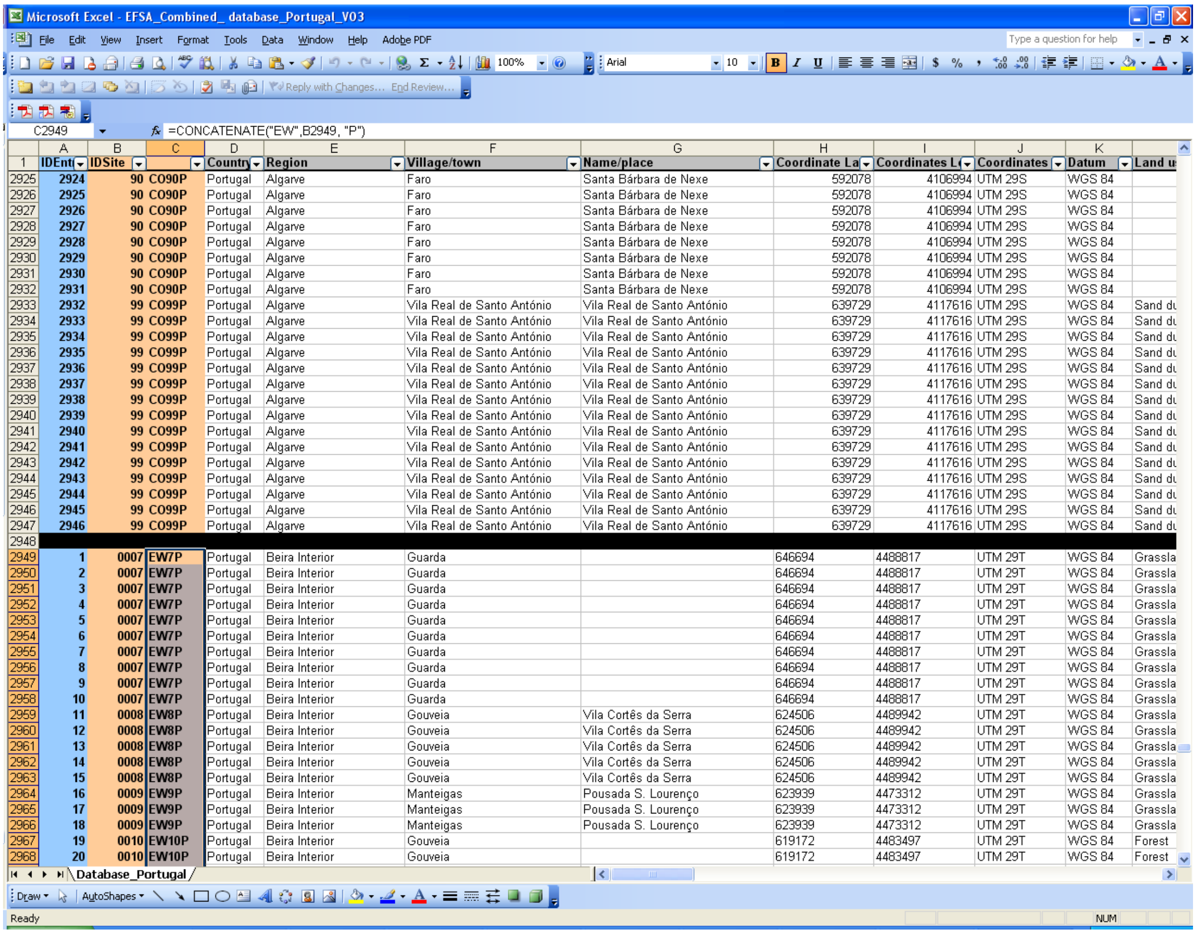
In order to project these data in the EU coordinate system (Lambert Azimuthal Equal Area), and to process the data in the most efficient way, it was necessary to reorganize the database:

- One global spreadsheet for each country was produced;
- From each of these global spreadsheets, partial spreadsheets were derived, grouping the records located in the same UTM zone
- In order to keep track of the changes, a new field was added (Figure 31), produced by the concatenation of:
  - Two capital letters for the organisms group (CO= collembola, EW= earthworms, IS= isopoda)
  - The numeric value of ID Site
  - The initial letter of the country name

These individual spreadsheets were exported in DB4 format, in order to be easily managed in ArcGIS. ArcGIS 9.3 is the GIS software that was used for the management and the analysis of the geographic information.

The following phase in the management of the data was the generation of Point Shapefiles, representing the locations in which the soil organism inventory was carried out and the re-projection of these maps.

The extraction of soil and climate data from the raster dataset, in correspondence of the soil organisms' survey points, was performed using the "Extract value to points" procedure; this procedure, that is a classical example of spatial query, allows the extraction of cell values of a raster, based on a set of points.



Microsoft Excel - EFSA_Combined_ database_Portugal_V03

C2949 =CONCATENATE("EW",B2949, "P")

	A	B	C	D	E	F	G	H	I	J	K	
1	IDEnt	IDSite		Country	Region	Village/town	Name/place	Coordinate La	Coordinates L	Coordinates	Datum	Land u
2925	2924	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2926	2925	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2927	2926	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2928	2927	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2929	2928	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2930	2929	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2931	2930	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2932	2931	90	CO90P	Portugal	Algarve	Faro	Santa Bárbara de Nexe	592078	4106994	UTM 29S	WGS 84	
2933	2932	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2934	2933	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2935	2934	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2936	2935	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2937	2936	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2938	2937	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2939	2938	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2940	2939	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2941	2940	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2942	2941	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2943	2942	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2944	2943	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2945	2944	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2946	2945	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2947	2946	99	CO99P	Portugal	Algarve	Vila Real de Santo António	Vila Real de Santo António	639729	4117616	UTM 29S	WGS 84	Sand du
2948												
2949	1	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2950	2	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2951	3	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2952	4	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2953	5	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2954	6	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2955	7	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2956	8	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2957	9	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2958	10	0007	EW7P	Portugal	Beira Interior	Guarda		646694	4488817	UTM 29T	WGS 84	Grassla
2959	11	0008	EW8P	Portugal	Beira Interior	Gouveia	Vila Cortês da Serra	624506	4489942	UTM 29T	WGS 84	Grassla
2960	12	0008	EW8P	Portugal	Beira Interior	Gouveia	Vila Cortês da Serra	624506	4489942	UTM 29T	WGS 84	Grassla
2961	13	0008	EW8P	Portugal	Beira Interior	Gouveia	Vila Cortês da Serra	624506	4489942	UTM 29T	WGS 84	Grassla
2962	14	0008	EW8P	Portugal	Beira Interior	Gouveia	Vila Cortês da Serra	624506	4489942	UTM 29T	WGS 84	Grassla
2963	15	0008	EW8P	Portugal	Beira Interior	Gouveia	Vila Cortês da Serra	624506	4489942	UTM 29T	WGS 84	Grassla
2964	16	0009	EW9P	Portugal	Beira Interior	Manteigas	Pousada S. Lourenço	623939	4473312	UTM 29T	WGS 84	Grassla
2965	17	0009	EW9P	Portugal	Beira Interior	Manteigas	Pousada S. Lourenço	623939	4473312	UTM 29T	WGS 84	Grassla
2966	18	0009	EW9P	Portugal	Beira Interior	Manteigas	Pousada S. Lourenço	623939	4473312	UTM 29T	WGS 84	Grassla
2967	19	0010	EW10P	Portugal	Beira Interior	Gouveia		619172	4483497	UTM 29T	WGS 84	Forest
2968	20	0010	EW10P	Portugal	Beira Interior	Gouveia		619172	4483497	UTM 29T	WGS 84	Forest

Database_Portugal

Figure 31: Structure of the country-based spreadsheet; the column with the new field has been outlined

D. Production of the Ecoregion Maps

The computation of the ecoregion maps has been based on the equations obtained in the data analysis (see Section 4.4), and was implemented using the Map Algebra tools of ArcGIS 9.3.1 (Raster Calculator, Single Output Map Algebra). In the Tables 17 and 18 the equations used for the computation of the earthworm and enchytraeids maps, respectively, are reported. The first set of equations, implying only the use of algebraic operators, has been calculated using the 'raster calculator", within the Spatial Analyst toolset, while the last expression, based on logical operators, was applied using the Single Output Map Algebra operator.

Table 17: Equations used for earthworms ecoregion maps

Map Algebra Operation with Raster Calculator

t_1	-0.498 + ([Cropland] * 0.0481) + ([Grassland] * 0.9844) +([Forest3] * -0.2298) + ([ph_top_efsa] * 0.317) + ([OC_efsa] * -0.0905) + ([tmean] * -0.2494) + ([Tdiff] * -0.0418)
t_2	2.7379 + ([Cropland] * -0.1215) + ([Grassland] * 0.2189) +([Forest3] * -1.1576) + ([ph_top_efsa] * 0.0567) + ([OC_efsa] * -0.0105) +([total_prec] * -0.0018) + ([tmean] * 0.0956) + ([Tdiff] * -0.1229)
z_1	Exp([t_1]) / (1 + Exp([t_1]))
z_2	Exp([t_2]) / (1 + Exp([t_2]))
ear_arr1	z_1
ear_arr2	[z2] * (1 - [z1])
ear_arr3	(1 - [z_2]) * (1 - [z_1])

Map Algebra Operation with Single Output Map Algebra

Earthworms Ecoregion Map	con(ear_arr1 >= 0.667, 1, ear_arr2 >= 0.667, 2, ear_arr3 >= 0.667, 3, arr1+arr2 >= 0.833 & ear_arr1 <= 0.667 & ear_arr2 <= 0.667, 12, arr1+arr3 >= 0.833 & ear_arr1 <= 0.667 & ear_arr3 <= 0.667, 13, arr2+arr3 >= 0.833 & ear_arr2 <= 0.667 & ear_arr3 <= 0.667, 23, arr1+arr2 <= 0.833 & arr2+arr3 <= 0.833 & arr1+arr3 <= 0.833, 123)

Table 18: Equations used for enchytraeids ecoregion maps

Map Algebra Operation with Raster Calculator

t_1	3.3243 + ([Grassland] * 0.4764) + ([Forest3] * 2.0354) + ([ph_top_efsa] * -0.2776) + ([OC_efsa] * -0.0206) + ([Clay] * -0.0114) + ([total_prec] * -0.0025) + ([tmean] * -0.2286) + ([Tdiff] * -0.0348)
t_2	-6.5979 + ([Grassland] * -0.5418) + ([Forest3] * 1.0585) + ([ph_top_efsa] * -0.2322) + ([OC_efsa] * -0.1102) + ([Clay] * -0.0505) + ([total_prec] * -0.0010) + ([tmean] * 0.3911) + ([Tdiff] * 0.2961)
z_1	Exp([t_1]) / (1 + Exp([t_1]))
z_2	Exp([t_2]) / (1 + Exp([t_2]))
enc_arr1	z_1
enc_arr2	[z2] * (1 - [z1])
enc_arr3	(1 - [z_2]) * (1 - [z_1])

Map Algebra Operation with Single Output Map Algebra

Enchytraeids Ecoregion Map	con(enc_arr1 >= 0.667, 1, enc_arr2 >= 0.667, 2, enc_arr3 >= 0.667, 3, arr1+arr2 >= 0.833 & enc_arr1 <= 0.667 & enc_arr2 <= 0.667, 12, arr1+arr3 >= 0.833 & enc_arr1 <= 0.667 & enc_arr3 <= 0.667, 13, arr2+arr3 >= 0.833 & enc_arr2 <= 0.667 & enc_arr3 <= 0.667, 23, arr1+arr2 <= 0.833 & arr2+arr3 <= 0.833 & arr1+arr3 <= 0.833, 123)

GLOSSARY AND ABBREVIATIONS

anecic earthworms	live in deep vertical burrows but feed at or near the soil surface and in the litter layer, especially at night;
BIOASSESS	EU founded research project; European Biodiversity Assessment Tools. EU FP5 project, Contract EVK4 -CT99-00280
Corine Land Cover 2000	EU Project under the CORINE programme (Coordination of information on the environment) aiming to produce cartographic information on soil use within Europe. See for more details Appendix B
EEC	European Economic Community
EFSA	European Food Safety Authority
endogeic earthworms	live and feed within the soil and burrow continuously to form a network of channels (mostly horizontal channels) in the first 20 cm of soil and around plant roots
ENVASSO	EU funded research project: ENVironmental ASsessment of Soil for mOnitoring, EU FP6, Contract No: 022713
epigeic earthworms	live in the litter layer (i.e. decaying plant debris, or compost);
EPPO	European Plant Protection Organisation
ERA	Environmental Risk Assessment
ERC	Ecotoxicologically Relevant Concentration
EU	European Union
FOCUS	FOrum for the Co-ordination of pesticide fate models and their USe
GD	Guidance Document
INSPIRE	Infrastructure for Spatial Information in Europe; The INSPIRE directive came into force on 15 May 2007 and will be implemented in various stages, with full implementation required by 2019. The INSPIRE directive aims to create a European Union (EU) spatial data infrastructure.
JRC	Joint Research Center (European Commission)
litter or 1 cm	Litter or 1 cm depth indicates that organisms are exposed via the litter layer, or if this is absent, via the upper layer of the soil (0 to 1 cm depth)
PEC	Predicted Environmental Concentration
PPP	Plant Protection Product
PPR	EFSA Panel / Unit on Plant Protection Products and their Residues
RA	Risk Assessment
Adjusted relative richness	Relative occurrence of species of a specific life form group per site. The percentage of the relative occurrence in relation to the sum of all trait groups defines the adjusted relative richness of that life form group on this site.
Raw relative richness	The percentage of a life form group in relation to the total number of

	species defines the raw relative richness of that life form group on this site.
SANCO	Directorate General for Health and Consumer Affairs (European Commission)
SGDBE	Soil Geographic Data Base of Europe
trait	A trait is a measurable property of organisms, usually measured at the species level and used comparatively across species. Examples of traits are structural traits (e.g. permeability of exoskeleton, lipid content, and complexity of the nervous system), morphological traits (e.g. size, volume / surface ratio), physiological traits (e.g. mode of respiration, detoxifying enzymes or digestive strategy), and ecological traits (e.g. mobility, feeding behaviour, trophic level, and place in the food web).
UTM	Universal Transverse of Mercator
VULCAN	EU founded research project: Vulnerability assessment of shrubland ecosystems in Europe under climatic changes. EU FP5 Contract EVK2-CT-2000-00094