Illuminating biodiversity changes in the ‘Black Box’

Helen R P Phillips†,§,¶, Erin K. Cameron‡, Nico Eisenhauer†,¶

† Saint Mary's University, Halifax, Canada
§ NIOO-KNAW, Wageningen, Netherlands
¶ German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany
¶ Leipzig University, Leipzig, Germany

Corresponding author: Helen R P Phillips (helen.phillips@smu.ca)

Received: 30 May 2022 | Published: 31 Aug 2022
Citation: Phillips HRP, Cameron EK, Eisenhauer N (2022) Illuminating biodiversity changes in the ‘Black Box’. Research Ideas and Outcomes 8: e87143. https://doi.org/10.3897/rio.8.e87143

Abstract

Soil is often described as a ‘black box’, as surprisingly little is known about the high levels of biodiversity that reside there. For aboveground organisms, we have good knowledge of the distribution of the species and how they might change under future human impacts. Yet despite the fact that soil organisms provide a wide variety of ecosystem functions, we have very limited knowledge of their distribution and how their diversity might change in the future. In order to create accurate and generalisable models of biodiversity, the underlying data need to be representative of the entire globe. Yet even with our recently compiled global earthworm dataset of over 11000 sites, there are gaps across large regions. These gaps are consistent across many other datasets of both above- and belowground diversity. In order to fill the gaps we propose a sampling network (SoilFaUNa), to create a comprehensive database of soil macrofauna diversity and soil functions (e.g. decomposition rates). Building on the existing dataset of earthworm diversity and early data from the SoilFaUNa project, we will investigate changes in earthworm diversity. From our current work, we know that both climate and land use are main drivers in predicting earthworm diversity, but both will change under future scenarios and may alter ecosystem functions. We will, using space-for-time substitution models, estimate how earthworm diversity and their functions might change in the future, modelling earthworm diversity as a function of climate, land use and soil properties and predicting based on future scenarios. Previous studies of aboveground diversity changes over time using time-series analysis...
have found no-net-loss in richness, but analyses have criticisms. We aim to use time-series data on earthworms to move this debate forward, by using data and statistical methods that would address the criticisms, whilst increasing our knowledge on this understudied soil group. Field experiments and micro-/mesocosm experiments have been used to investigate the link between a number of soil organisms and ecosystem functions under few environmental conditions. Meta-analyses, which can produce generalisable results can only answer questions for which there are data. Thus, we have been lacking on information on the link between the entire community of soil fauna and ecosystem functions and impact of changes to the soil fauna community across environmental contexts. Using data collected from the SoilFaUNa project, we will, for the first time, synthesise globally distributed specifically-sampled data to model how changes in the community composition of soil macrofauna (due to changes in land use, climate or soil properties) impact the ecosystem functions in the soil.

**Keywords**

earthworms, soil macrofauna, soil functions, ecosystem functions, global biodiversity monitoring

**List of participants**

Dr. George Brown, Dr. Ika Djukic, Dr. Carlos Guerra, Dr. Patrick Lavelle

**State of the art and preliminary work**

**Loss of biodiversity**

The diversity of life on Earth is declining through species extinctions (Butchart et al. 2010, Dirzo et al. 2014) and reductions in population abundances (Butchart et al. 2010, Collen et al. 2009). Current species extinction rates are estimated to be orders of magnitude higher than the background rate in geological history from the fossil record (Pimm et al. 2014). Land-use change is one of the main drivers of this loss (Sala et al. 2000, Pereira et al. 2010) and unless measures are taken to manage these changes, rapid biodiversity loss is predicted to continue (Pereira et al. 2010, Tittensor et al. 2014). Biodiversity is also changing at local scales (Vellend et al. 2013, Dornelas et al. 2014, Phillips et al. 2017a, Blowes et al. 2019), but the magnitude and direction of this change is still being debated (Gonzalez et al. 2016, Vellend et al. 2017, Cardinale et al. 2018). Investigating biodiversity changes at the local scale is important as local biodiversity can be linked directly to ecosystem services and functioning (Cardinale et al. 2012, Hooper et al. 2012, Eisenhauer 2019).

Recent large-scale synthesis analyses have shown how local biodiversity (i.e. a measure of the ecological assemblage within a sampled plot, which can vary in size from metres to kilometres, *sensu Newbold et al. 2015*) is being lost as a result of anthropogenic
disturbance, in particular land-use change (Newbold et al. 2015, Newbold et al. 2016, Phillips et al. 2017a) and habitat loss (Phillips et al. 2018). However, the data used in these large-scale synthesis analyses (Dornelas et al. 2014, Newbold et al. 2015) often have biases, both geographically and taxonomically (Phillips et al. 2017b, Cameron et al. 2018) and, in particular, are often heavily biased towards certain taxa, mainly plants, birds and some marine organisms. Soil biodiversity is very rarely considered in such large-scale synthesis (Phillips et al. 2017b), despite its importance to ecosystem functions and services (Bardgett and van der Putten 2014, Wall et al. 2015). Thus, soil is still regarded as a ‘black box’ in the context of biodiversity change (Phillips et al. 2017b). The amalgamation of data from a wide-range of taxa also means that the often-hypothesised link between biodiversity loss and ecosystem services cannot be tested, as, for the majority of taxa included, there is no quantitative relationship. However, focussing on single taxa groups where there is a known link to an ecosystem function or service they provide (e.g. earthworms and decomposition; Coleman et al. 2004) allows the relationship between changing diversity and the provision of ecosystem functions to be quantitatively estimated.

**Earthworms as a model organism for biodiversity research**

Organisms that are ecosystem engineers, those that create, modify or maintain habitats (Jones et al. 1994), such as earthworms (Lavelle et al. 1997), can have especially large impacts on ecosystem functions. Earthworms provide a variety of ecosystem functions and services that are critical for human well-being (Blouin et al. 2013), such as increasing crop production and aboveground biomass (van Groenigen et al. 2014), as well as regulating the carbon sink of the soil (Zhang et al. 2013). They can be categorised into three main ecological functional groups; litter dwellers (epigeics), soil feeders (endogeics) and deep-burrowers (anecics) (Bouché 1977). The contribution to the provisioning of ecosystem functions and services varies depending on the functional group of the earthworm species (Brown 1995, Eisenhauer 2010, Lubbers et al. 2013, Craven et al. 2017). Epigeic species are typically found in the upper layers of the soil and litter and are important in the first stages of physical breakdown of the litter layer (Brown 1995), increasing the surface area of the litter for microbial decomposition (Hättenschwiler et al. 2005). Endogeic species live in the upper mineral soil layers, creating horizontal burrows (Bouché 1977, Brown 1995), with most species creating both above- and belowground casts. These casts alter the structure (porosity and aggregation) of the soil, changing the water infiltration, run-off and water holding capacity (Ernst et al. 2009). Anecic species are the deep burrowing species, moving litter from the surface into the deeper layers of the soil, whilst moving mineral soil from the depths to the surface via their cast production (Edwards and Bohlen 1996).

The relatively large body size of earthworms compared to other soil taxa (Veresoglou et al. 2015) means that they are, to some degree, easy to sample and identify to species level and, thus, often more studied than other soil taxa. Yet, despite the large amounts of data (e.g. Phillips et al. 2019 recently collated nearly 200 datasets including > 11,000 sampling locations from across the globe), surprisingly little is known about the diversity of earthworms across large regions or globally (Hendrix et al. 2008, Cameron et al. 2016). Whereas many global distribution maps and diversity maps have been produced using aboveground taxa (e.g. Orme et al. 2005, Kreft and Jetz 2007, Roll et al. 2017),
biogeographic studies of earthworms have only occurred across smaller regions (Rutgers et al. 2016) or at coarse resolutions (either spatially or taxonomically; Hendrix et al. 2008). In order to fill this gap, we are currently analysing spatial patterns of earthworm communities across the globe, using our large dataset of sampled earthworm diversity (Phillips et al. 2019).

The work that we are currently undertaking will further our understanding of the current distribution of earthworm diversity and the importance of different environmental drivers in shaping the communities (Phillips et al. 2019). However, we know that most environmental drivers will change as a result of anthropogenic impacts (Pereira et al. 2010). For example, future predictions indicate that agricultural land will increase as a result of the increasing human population (Lambin and Meyfroidt 2011) and precipitation will change as a result of climate change (IPCC 2014). Due to the importance of earthworms for ecosystem services, we need to fully understand how earthworm diversity might change under these future scenarios of change and how the ecosystem functions they provide might be altered in response. However, until data gaps have been filled, confidence in future projections of earthworm diversity will be low and our ability to answer important questions across large scales will be substantially limited.

**Gaps in global datasets**

For decades, papers have commented on the lack of biogeographic studies for soil fauna (Brussaard 1997, Rusek 1998, Eisenhauer et al. 2017). Yet despite the repeated calls, only now are biogeographic studies of soil organisms starting to appear on specific taxa (Tedersoo et al. 2014, Rutgers et al. 2016, Delgado-Baquerizo et al. 2018, Ramirez et al. 2018). One of the reasons why biogeographic studies may not have appeared earlier (unlike those of aboveground organisms, such as birds: Orme et al. 2005, plants: Kreft and Jetz 2007 and reptiles and amphibians: Roll et al. 2017), may be due to soil ecologists' primary focus on local-scale research and aboveground macro-ecologists' underestimation of the amount of data available. Thus, one of the ways to facilitate the inclusion of soil biodiversity data into macro-ecological studies would be the creation of a soil biodiversity database (Ramirez et al. 2015, Phillips et al. 2017b), which has already been done for many individual groups of taxa globally (e.g. ants: Dunn et al. 2007, plants: Kreft and Jetz 2007, bacteria: Delgado-Baquerizo et al. 2018) or for regions (e.g. Edaphobase for earthworms in Germany: Burkhardt et al. 2014). Increasing the ease of access to soil biodiversity data will likely result in these taxa being included in further analyses. For example, the TRY database, which collates functional traits of plants from across the globe Kattge et al. 2011, has now been used in over 150 further publications (https://www.try-db.org).

However, in all of the globally assembled databases (e.g. PREDICTS: Hudson et al. 2017, BioTIME: Dornelas et al. 2014, the Global Ant Database: Dunn et al. 2007 and the global databases of bacteria: Delgado-Baquerizo et al. 2018, Ramirez et al. 2018), there are regional gaps in data availability, with large regions of the tropics and Russia represented by minimal data. It is possible that limited sampling has occurred in these regions and it is highly likely that any data generated from these regions is being published in languages
other than English (Amano et al. 2016). These gaps will result in at least two issues; firstly, geographic biases in the underlying data will increase the risk that results are not transferable across the entirety of the globe, especially if responses of biodiversity vary across different regions (Phillips et al. 2017a). Secondly, if the data do not encompass a large enough disturbance gradient, any changes in response to that disturbance may be under- or over-estimated (Elahi et al. 2015). For example, analysis of the sampled regions from the largest collation of earthworm data (Phillips et al. 2019), show how the 11,009 sites do not encompass areas that will experience some of the largest changes in climate, despite being within travelling distances of major settlements (Fig. 1a). Failing to capture major threats faced by earthworms, or any other organism, may result in inaccurate estimates of how biodiversity changes under these threats.

Figure 1. A) Black dots show sample locations of the 11,009 sites in the Phillips et al. 2019 data. Yellow indicates areas where climate change is predicted to be greatest (> 3 degrees C change in temperature or a change of absolute 100 mm in rainfall; using Karger et al. 2017 and are relatively accessible from settlements (< 12 hours travel time; using Weiss et al. 2018). Apart from western Europe, most regions with a lot of climate change are not sampled representatively. B) Environmental representation of the current database across 10 environmental variables. Coverage is typically well spread, but under-representing certain values.
Such gaps in global datasets could be filled by developing a global network of researchers who apply standardised protocols across sites. Previously, standardised networks of experiments have been useful in addressing typically local-scale ecological questions across a global scale (e.g. NutNet: Borer et al. 2014, TreeDivNet: Verheyen et al. 2016, dummy caterpillars: Roslin et al. 2017). A standardised protocol ensures that data are easily comparable, with similar assumptions and biases (Borer et al. 2014). Furthermore, networks are also able to address questions not possible using meta-analytical approaches due to a lack of data in the primary literature (Fraser et al. 2013). Such schemes have been successful in the past, answering questions, such as the relationship between plant diversity and productivity (Borer et al. 2017), as well as creating new collaborative networks of researchers. For example, the TeaComposition project (Djukic et al. 2018) (based on the methodology of the Tea Bag Index; Keuskamp et al. 2013) enlisted over 200 researchers from across the globe in just over a year. Within such experimental as well as monitoring networks, recruitment of researchers can be targeted to certain areas to ensure the best possible geographical coverage of sites, thereby reducing geographic biases and resulting in more accurate models (Maestre and Eisenhauer 2019).

**Projecting changes in earthworm diversity under future scenarios of change**

With predicted future changes in anthropogenic pressures, such as climate change (IPCC 2014) and land use change (Sala et al. 2000, Lambin and Meyfroidt 2011), it is important that we understand how local biodiversity might respond and how this might impact ecosystem functions and services. For example, Newbold et al. 2015 provided some of the first global projections of changes in local biodiversity due to future land use change (with changes in land use calculated from projections of climate change and their mitigation scenarios; Hurtt et al. 2011). They showed that under the ‘Business-As-Usual’ scenario (MESSAGE8.5), local biodiversity is predicted to decline by an additional 7% by 2100 compared to estimates of the current day biodiversity levels.

There are few studies that use future scenarios to project changes in local biodiversity at global scales. Those that do are often limited to only one driver of change, for example, land use (Newbold et al. 2015) or climate (García Molinos et al. 2016). However, impacts of climate and land use are likely to have interactive effects on biodiversity (Frishkoff et al. 2016). None of the studies has included any soil biodiversity data (Phillips et al. 2017b) and there has been no specific assessment of the effect of future changes on soil biodiversity, even for more well-studied taxa, such as earthworms. Therefore, it is unclear whether earthworms may respond particularly negatively to the conditions that may be present in these scenarios, despite their importance for ecosystem function and services (Blouin et al. 2013, van Groenigen et al. 2014).

**Time series analysis**

There is an ongoing debate surrounding the direction and magnitude of local biodiversity change, with some synthesis studies showing loss of local biodiversity (Newbold et al. 2015, Phillips et al. 2017a) and others showing ‘no-net-loss’ of diversity (Vellend et al. 2013, Dornelas et al. 2014, Hillebrand et al. 2018), but rather changes in community...
composition and structure (Dornelas et al. 2014, Hillebrand et al. 2018). However, the modelling approaches in these studies are very different. The ‘no-net-loss’ studies used time-series analysis (using data from sites that have been sampled more than once across multiple months or years), while the other studies used ‘space-for-time’ substitution (which assumes that spatially distributed sites differing in their disturbance will have the same biodiversity difference as sites that change their disturbance over time; e.g. Newbold et al. 2015, Phillips et al. 2017a). It is highly likely that these differing methodologies have led to the differences in the conclusions (De Palma et al. 2018).

Although analysis of time-series data is highly valuable for examining dynamic changes in local biodiversity, there have been criticisms of both the Vellend et al. 2013 and Dornelas et al. 2014 studies (Cardinale 2014, Gonzalez et al. 2016, Cardinale et al. 2018). These criticisms have focussed on several aspects of the modelling approach: 1) Spatial biases within the datasets used, with species-rich areas and areas under pressure from anthropogenic impacts being under-represented (Gonzalez et al. 2016). 2) Analysis of data without a reference (e.g. an undisturbed baseline; Eisenhauer et al. 2016). It is assumed that the start of the time-series contains diversity estimates similar to those found in the pristine conditions, which may not be true. This could explain the wide range of responses seen across the datasets, as over time, the diversity at some time-series could be recovering, whilst other time-series could be facing worsening pressures (Eisenhauer et al. 2016). 3) Investigating changes in temporal diversity outside of the context of external pressures, especially anthropogenic pressures (Cardinale 2014, Elahi et al. 2015, Cardinale et al. 2018). This may result in the average trends reported not being representative of biodiversity change across ecosystems globally (Elahi et al. 2015).

The synthesis of space-for-time data is often used as more data are available, as primary datasets require only one season of fieldwork (De Palma et al. 2018). Another advantage is that it is easy to relate the site level data to simultaneously occurring external pressures, such as land use or climate. Obtaining accurate data across multiple years can be problematic. However, there are criticisms of this approach. The underlying assumption is that spatial comparisons are a suitable substitute for temporal changes, which often is not true. Indeed, space-for-time studies may consistently underestimate biodiversity loss (França et al. 2016). Another issue with this approach is the inability to show dynamic changes in biodiversity, for example, recovery after a disturbance (Dunn 2004) or an extinction debt (Vellend et al. 2006).

Both techniques have downfalls and, consequently, an ideal solution is to combine the two approaches. Using time-series data to show dynamically how local biodiversity is changing, while addressing previous criticisms by incorporating baselines and information on external anthropogenic pressures, may move this debate forward. Given the amount of earthworm data available, and the fact that a time-series analysis has not been performed on any soil taxa, earthworms are an ideal study organism with which to combine these approaches.
**Loss of ecosystem services**

Soil biodiversity is a key driver of many vital ecosystem functions (Bardgett and van der Putten 2014). However, much of our understanding of the quantitative relationships between soil diversity and ecosystem functions comes from small-scale field studies or micro-/mesocosm experiments (Bradford et al. 2002, Heemsbergen et al. 2004, Wagg et al. 2014). Field (experimental and observational) studies typically focus on a single taxa group or trophic level (e.g. Eisenhauer et al. 2009) and ignore other groups. In contrast, micro-/mesocosm experiments often examine the effects of multiple taxonomic and functional groups on ecosystem functions (such as productivity; Bradford et al. 2002 and decomposition; Heemsbergen et al. 2004) and attempt to isolate the effects of individual taxa (Wagg et al. 2014). However, micro-/mesocosm studies may be unrealistic, as they often involve species-devoid communities, with manipulation treatments that are unlikely to occur in natural settings (Bradford et al. 2002). Therefore, results may not be generalisable from micro-/mesocosms and single, small-scale field studies to large scales.

As an alternative, meta-analysis can use globally distributed datasets to create more generalisable results (Koricheva et al. 2013). A disadvantage is that meta-analyses are limited by the number of available primary studies that investigate a similar question (Koricheva et al. 2013). This can often result in small sample sizes, whilst limiting the questions being asked. Previous meta-analyses have shown how earthworms can influence ecosystem services, such as increasing crop production and aboveground biomass (van Groenigen et al. 2014). However, due to the nature of the primary literature, these meta-analyses assume that changes in ecosystem functions are only caused by the presence or absence of earthworms and not other soil taxa. Moving the field forward may rely on establishing networks of researchers working towards common questions. For example, Wall et al. 2008 and de Vries et al. 2013 developed networks of researchers to investigate specific aspects of how soil community composition contributes to ecosystem functions.

Disturbances (a term used here to describe a site that differs from a ‘reference’ site, for example, a change in land use or soil properties and, therefore, not necessarily due to human impacts) are likely to change the composition of communities, but exactly how composition will change may vary. Groups of taxa may be lost or changes in abundance or biomass may occur (Hillebrand et al. 2008). These changes in the composition of the community will likely impact the level of the ecosystem functions provided: 1) As a reference, in an undisturbed area, the composition of the community supports a given level of ecosystem function (Fig. 2a; Fox and Kerr 2012). 2) Following a disturbance, all groups of taxa may decrease equally in abundance, reducing the total abundance of the community. Therefore, we might expect ecosystem function to decline (Fig. 2b; Fernández et al. 2015). 3) However, some taxa groups in the community may increase in abundance, rather than decrease. Thus, total community abundance remains equal, but the composition substantially changes (Cesarz et al. 2017). This could result in a reduction in ecosystem function if the dominant taxa group is unable to compensate for the lost abundance of the other taxa groups (Fig. 2c; Hunt and Wall 2002). 4) Alternatively, if the dominant taxa group can maintain the ecosystem function, no reduction would occur.
Despite its importance, very little is known about the biodiversity in the ‘black box’ (Phillips et al. 2017b). Although steps have been taken in attempts to show biogeographic patterns of some taxa, there are still gaps in our knowledge of how biodiversity might be changing due to the changing environment (Cameron et al. 2018). In addition, how changes in biodiversity might impact the provision of ecosystem functions that we heavily rely upon is unclear.
Project-related publications

Articles published by outlets with scientific quality assurance, book publications and works accepted for publication, but not yet published.


Objectives, concept and approach

Objectives

Although there have been many small-scale studies on changes in soil biodiversity, there has been little progress in creating large-scale, generalisable results. This project aims to investigate changes in soil biodiversity (with a focus on earthworms, Order:
Crassiclitellata), across large spatial scales using synthesis analyses. In addition, the project will also link soil biodiversity to soil functions and investigate how any changes in biodiversity might impact the ecosystem functions upon which we rely. Using previously collated data, as well as data collected specifically for this project and appropriate statistical methods that deal with the complexities of the data and the ecological questions, we will advance the field of global soil biodiversity.

Although soil biodiversity data are available across the globe the distribution is poor, especially within certain regions, and heavily biased towards certain realms and environmental conditions (Fig. 1). In order for any biodiversity model to be as transferable and accurate as possible, data need to be as representative of as much of the terrestrial realm, and all its environmental gradients, as possible. Therefore, we will be creating a network of researchers to collect data on soil biodiversity and soil ecosystem functions in a standardised, simple and cost-effective way (WP1). This collected data will be used in WP4, and if possible, WP2 and WP3.

WP2, 3 and 4 aim to further both our knowledge of the changes in soil biodiversity, particularly earthworms, whilst also furthering key, timely and highly relevant questions in the field of biodiversity change and ecosystem function. Previous synthesis analyses have shown how biodiversity is predicted to respond under future scenarios of change and have investigated whether biodiversity has been changing over time. However, these questions have been studied using datasets primarily composed of aboveground organisms. WP2 and 3 will modify the previously-used methods, to answer these questions in relation to earthworms. In addition, by modifying the methods previously used, we aim to further the field by addressing key questions that have not been addressed previously, such as whether local earthworm biodiversity is changing over time.

Providing this research is crucial, given that the understudied soil biodiversity is relied upon heavily for many ecosystem functions. In WP2, we aim to show how changes in earthworm diversity (as predicted based on future scenarios of global change) might affect the ecosystem functions that they provide. While, in WP4, we will further extend the biodiversity-ecosystem function field, by researching how multiple ecosystem functions provided by the soil might change when the soil community is altered, using specifically-collected datasets from non-manipulated biodiversity measurements in a globally-distributed network.

This project will be led by Dr Helen Phillips in the Experimental Interaction Ecology group of Prof. Nico Eisenhauer, at the University of Leipzig and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig. Dr Helen Phillips was the lead post-doc (as part of a working group led by Prof. Nico Eisenhauer) collating data of earthworm diversity measures. This dataset now allows for a number of novel analyses.

**Anticipated total duration of the project**

Current funding: none.
Duration: 3 years

Work programme including proposed research methods

WP1 - Monitoring Scheme

Objectives

The number of research papers using synthesis approaches (i.e. collating raw data from previously published papers to conduct new analyses) to address patterns in biodiversity and questions on biodiversity change has recently increased (Dunn et al. 2007, Dornelas et al. 2014, Newbold et al. 2015, Phillips et al. 2017a, Delgado-Baquerizo et al. 2018, Ramirez et al. 2018, van den Hoogen et al. 2019). However, most have consistent gaps in where datasets are located (see Dunn et al. 2007, Dornelas et al. 2014, Hudson et al. 2017) -- usually lacking data from central Africa, boreal Asia and, to some extent, South America (Cameron et al. 2018). In addition, the datasets included in syntheses may not be equally distributed across the gradient of threats to biodiversity (Gonzalez et al. 2016) (Fig. 1A) or environmental conditions (Fig. 1B). For example, studies may be less common in regions that are facing large amounts of climate change (Fig. 1A). Until we have representative data, across regions and threats, our understanding of biodiversity change will remain limited.

We will create a network of researchers (the Soil biodiversity and Soil Function Network-SoilFaUNa) to collect soil biodiversity and soil functions data from regions that are typically understudied. The protocols used will be standardised and simple, where possible being based on protocols developed and agreed upon by soil experts within iSBio (https://home.uni-leipzig.de/div/isbio/; Heintz-Buschart et al. 2020). In addition, the protocol is cost-effective, enabling implementation with few resources (such as in the TeaComposition initiative; Djukic et al. 2018) in order to include geographic areas not included in previously established global sampling and monitoring networks. We will initially develop the network by re-contacting earthworm researchers we have previously worked with (~ 200 researchers, Phillips et al. 2019) to invite them to collect additional data. In a survey we previously conducted of 77 earthworm researchers, 89% indicated they would be willing to collect additional data for a global database. To find additional participants, we will locate universities within target regions and approach researchers with some expertise in soil ecology. Papers advertising the creation of the SoilFaUNa project will also be published at the start of the project, in continent-specific journals (e.g. African Journal of Ecology), through organisations (e.g. Global Soil Biodiversity Initiative) and other networks (e.g. TeaComposition network, ~ 200 researchers; Djukic et al. 2018), so that researchers can also get in contact with us. Data from this WP will feed into analyses in WP2, WP3 and WP4.

Methods

At the start of the project, we will re-contact the ~ 200 earthworm researchers we have previously worked with. We will ask researchers who have collected earthworm data in the
past if they can resample the earthworm communities of their previously-sampled sites using their original methodology, in order to obtain more time-series datasets (suitable for analysis in WP3). They will also be informed about our new standardised sampling protocol (see below) and will be invited to start new sampling campaigns to collect data suitable for WP2 and WP4 (soil fauna and functions). We will also locate new soil ecologists to collaborate with, by searching through universities in target regions, for example across the tropics (i.e. Indonesia), central Africa (i.e. Tanzania), as well as boreal Asia (i.e. northern China). We believe that our advertising papers will also prompt soil ecologists to contact us. Nico Eisenhauer has successfully tested this approach recently in a global collection of soil microbial biomass data (Smith et al. 2020). Researchers who have not previously collected earthworm data will only be invited to use our standardised protocols.

To standardise sampling and assist researchers who may have limited experience in sampling soil fauna, an easily reproducible protocol has been drafted and will be distributed to all collaborating researchers (Suppl. material 1; SoilFaUNa Sampling Protocol). The protocol is based on methods suitable for both tropical and temperate regions (Anderson and Ingram 1993, ISO 2018) and will be tested and refined by the post doc and the international collaborators prior to be distributed to collaborating researchers. The remainder of this methods section refers only to our standardised protocols. To begin with, we are aiming for sampling at 200 new sites with our standardised protocols, which is manageable (in terms of lab processing) and achievable in the appropriate timeframes.

Collaborating researchers for the 200 new sites will be asked to sample earthworms, other soil macrofauna and soil functions at more than one site at least twice (at months 3 and 12 when teabags are collected; see below). Within each collaborating researcher’s sampling campaign, sites must be distributed across a range of current environmental conditions (e.g. land use, habitat cover, soil, climate), in the hope of increasing the environmental coverage of the database as a whole (Fig. 1b), and some information must be available on the history of the sites. Earthworm and soil macrofauna sampling will follow accepted protocols (Anderson and Ingram 1993, ISO 2018), using a handsorting approach (a block of 50 cm x 50 cm x 25 cm deep is excavated and soil macrofauna removed) followed by application of mustard solution into the created hole (full SoilFaUNa sampling protocol in Suppl. material 1). This combined approach has been shown to be the most accurate for obtaining a representative sample of earthworm communities at a site (Lawrence and Bowers 2002), whilst handsorting is also suitable for sampling a wide-range of soil macrofauna (ISO 2018). For each sample, the total abundance and fresh biomass of each soil fauna group (e.g. earthworms, carabid beetles, centipedes, millipedes etc.) will be measured. Counting and weighing organisms within fauna groups is similar to the approach used by the MacroFauna project (led by Patrick Lavelle), which successfully sampled soil macrofauna at over 2000 sites across the tropics, Europe and Australia (Orgiazzi et al. 2016, Cameron et al. 2018). Even when not identified to species level, earthworm data can still be incredibly useful, as density and biomass are critical indices (Bouché and Al-Addan 1997, Winfree et al. 2015) that will be useful to achieve our goals and advance the field. As further analysis will be focussed on earthworms (WP2, WP3, WP4), if the researcher has the experience (or access to region-specific taxonomic keys),
Earthworms will be identified to species or morphospecies. Where necessary, to help with differentiation of morpho-species, we hope to also create and distribute a simplified key of external characteristics (e.g. pigmentation, body size, clitellum type, setal pattern etc.) to collaborating researchers. The biomass and abundance of each earthworm (morpho-) species will then be calculated for each site, in addition to the total biomass and abundance.

Soil functions will also be measured at each site. To measure the belowground decomposition rate, the Tea Bag Index method will be used, which estimates decomposition rates by calculating weight differences in tea bags over time (see Keuskamp et al. 2013 for full protocol). In short, eight tea bags (4 x Lipton Green tea bags and 4 x Lipton Rooibos tea bags; provided by UNILEVER to ensure consistency) will be buried within the sampling site at the bottom the H horizon. The tea bags will be checked at 3, 12, 24 and 36 months. We will also ask for soil samples from each site to be sent to iDiv for analysis (100 g fresh weight). From each soil sample, we will measure soil pH, aggregate stability, microbial C, microbial respiration, metabolic quotient (qO2) (Eisenhauer et al. 2018), gravimetric soil water content and carbon:nitrogen ratio in the Eisenhauer laboratory that has a permit for processing and storing international soil samples. Following the iSBio protocols, we also hope to ensure consistency across samples by having the samples frozen prior to shipping (likely thawing during and post-flight) and by storing the samples for the same length of time between shipping and processing. This approach has been proven successful in previous global soil collections by the Eisenhauer lab (e.g. in Nutrient Network: Risch et al. 2019, in TreeDivNet: Cesarz et al. 2020).

All collaborating researchers will be sent a standardised data template, into which they will enter meta-data and biodiversity data and soil function data from each site sampled. This will ensure that all necessary data are collected and are stored in a standardised form, thereby decreasing the time needed to process it ready for analysis. In addition to our own global initiatives (e.g. Risch et al. 2019, Maestre and Eisenhauer 2019, Heintz-Buschart et al. 2020, Cesarz et al. 2020), we will build on the experience of Ika Djukic, who has coordinated the global TeaComposition project (Djukic et al. 2018) and has ample experience with obtaining relevant local data in a standardised way.

Mustard powder, teabags and containers for soil samples will be sent to all collaborating researchers, with the remaining costs of all fieldwork being covered by the collaborating researcher; however, protocols have been designed to use inexpensive methods/equipment. Collaborating researchers will be responsible for acquiring any permits needed for collecting and moving soil samples out of their country. We will obtain any permits that are needed to move soil samples into Germany (such as done in previous international, collaborative projects, e.g. Heintz-Buschart et al. 2020, Cesarz et al. 2020). We are aware that moving samples out of certain countries will not be possible. In these circumstances, we will work closely with and provide funding to soil scientists within the country, who would undertake all soil and biological processing. For example, Brazil has strong regulations regarding movement of material, but we have previously collaborated with Prof. Ademir Araújo who has agreed to undertake the processing when necessary, as done in previous collaborations (e.g. Araújo et al. 2014).
Data that are analysed at iDiv would be shared with the original collaborating researcher within the timeframe of the project. A subsample of 2 g of all soil samples would be stored at -80°C long term (at iDiv) to enable genetic analyses to be undertaken in the future. In addition, all biodiversity data would be entered into the Global Soil Biodiversity Database. We will ask all collaborating researchers to store biological samples for five years post sampling in 95% alcohol (see sampling protocol in Suppl. material 1). Stored samples will need to remain accessible, but can be at a place of the collaborating researchers choosing (i.e. in their lab collections or at local museums). Information about storage locations of reference samples will also enter the database. Collaborating researchers will be encouraged to publish analysis from their own biodiversity data, after they have collected and shared it with the SoilFaUNa project. In addition, they will be able to publish products using the analysed soil samples, once the data have been returned to them. In addition to being an author on other manuscripts when their data have been used in other WPs.

2.3.2 WP2 - Projecting changes in earthworm diversity under future scenarios of change

Objectives

From previous work, we are beginning to understand the large-scale spatial patterns of earthworm diversity in relation to current environmental conditions, such as soil properties, land use and climate (Phillips et al. 2019). However, human impacts are changing current conditions and, thus, the future state of many ecosystems will be altered (IPCC 2014, IPBES et al. 2016). It is vital that we understand how biodiversity will respond (Newbold et al. 2015) and especially how organisms that we rely upon for many ecosystem functions and services respond (e.g. earthworms, Blouin et al. 2013). Yet, we know surprisingly little about how soil organisms may respond to changing environmental conditions (Bardgett and van der Putten 2014).

We propose, using previously collected data (11,009 sites, Phillips et al. 2019), as well as data collected as part of WP1, to model how earthworm diversity may change in response to changing environmental variables using a space-for-time approach. Previous work has shown that earthworm diversity is low in agricultural land uses (over 20% reduction in species richness compared to the highest diversity land uses (Phillips et al. 2019) and with higher annual mean temperatures, both of which are projected to increase under future scenarios of change (Lambin and Meyfroidt 2011, IPCC 2014). Therefore, we hypothesise that local earthworm diversity will be reduced under these future scenarios of human impacts (WP2-H1). As earthworm functional groups vary in their contribution to ecosystem functions (Lubbers et al. 2013, van Groenigen et al. 2014) and may vary in their responses to a changing environment, we will also project how changes in earthworm community composition, with respect to functional group, may impact ecosystem functions.
Methods

For this WP, we will use previously-collected data (Fig. 1A, Phillips et al. 2019), as well as any data available from the early stages of WP1. We will also undertake a literature search and collate raw data from suitable recent studies. Once the dataset has been compiled, we will create models that predict earthworm diversity (biomass, abundance and diversity) using environmental variables and project how earthworm diversity may change under future scenarios of change.

For datasets to be suitable for this analysis, they will need to contain earthworm data from across more than two sites where current environmental conditions (e.g. land use, climate, soil properties) vary. The exact position of each site would also need to be known (from either GPS coordinates or by digitising and geolocation of available maps).

In order to create biodiversity projections using scenarios of future change (e.g. Hurtt et al. 2011, Riahi et al. 2017), mixed effects models will be constructed (using ‘lme4’ in R; Bates et al. 2015, R Core Team 2016) containing predictor variables encompassing environmental conditions considered in the future projections (e.g. temperature), but also environmental conditions that are important for earthworm diversity that have not been projected into the future (e.g. soil properties). For each site, some predictor variables will come from the original data collector (e.g. land use, with site descriptions classified into categories, based on the Representative Concentration Pathways harmonised land-uses; Hurtt et al. 2011) or from matching site coordinates to freely-available global data layers (e.g. annual mean temperature, available from CHELSA Climate; Karger et al. 2017). Models will be created for the three biodiversity metrics, species richness (the most commonly reported biodiversity measure), abundance (a metric increasingly being used in biodiversity studies; Winfree et al. 2015, Newbold et al. 2016) and biomass (which in earthworm communities can often be linked to ecosystem function: Bouché and Al-Addan 1997), as well as for the three main functional groups of earthworms (epigeics, endogeics and anecics). The models will then be used to predict earthworm diversity under future scenarios of change (e.g. RCPs, SSPs; Hurtt et al. 2011, Riahi et al. 2017) using variables from the global layers of future environmental conditions. In addition, we will use previously-published relationships between earthworm diversity and ecosystem functions (e.g. aboveground productivity; van Groenigen et al. 2014) to predict how the ecosystem function will change given the modelled reduction in earthworm diversity. This analysis will help us identify regions that may be particularly vulnerable to earthworm-induced change in ecosystem function, so they can receive more scientific attention and could be the focus of future monitoring campaigns.

2.3.3 WP3 - Time-series analysis

Objectives

Debate has continued over whether local biodiversity is declining (Gonzalez et al. 2016, Vellend et al. 2017, Cardinale et al. 2018). The debate has focused around two methodologies -- space-for-time and time-series analysis. However, both are not without
their flaws (França et al. 2016, De Palma et al. 2018). To determine if local biodiversity is naturally changing over time or only as a direct result of human impacts and changing environmental conditions, long-term data are needed. Earthworms are an ideal study organism to help progress this debate. Although there are little data available on changes in soil biodiversity over time, some time-series data on earthworm biodiversity are available (currently 17 datasets compiled as part of Phillips et al. 2019, which encompass 457 sites, across 12 countries) and, due to being understudied, there is no prior expectation about how their diversity is changing.

In this WP3, we will use the time-series analysis approach similar to that used by Dornelas et al. 2014 but will incorporate changes into the statistical analysis proposed by Gonzalez et al. 2016 and De Palma et al. 2018. Specifically, we will investigate if local earthworm diversity (biomass, abundance and diversity) has changed over time and whether the response is consistent when land use (and other global change drivers) are accounted for. We hypothesise that any change in biodiversity will have been caused by environmental conditions and human impacts, and time will not explain any additional variation (WP3-H1).

Methods

Suitable datasets collected previously (Phillips et al. 2019) will be used in this analysis. In addition, a literature search will be undertaken to collate recent datasets. In order for data to be suitable for a time-series analysis, each dataset needs to contain multiple sites (which vary in their disturbance) that were sampled for earthworms on at least two occasions across different years, using a consistent methodology. As part of WP1, we will have contacted earthworm researchers asking them to resample their sites using their previous methodology. Any data from resampling will also be used in this analysis.

The earthworm samples from each sampled time will be temporally matched to both soil properties (measured either by researchers at the same time as sampling or from global data layers, such as SoilGrids; Hengl et al. 2017) and climate variables. Land use/habitat cover will be classified at each site, as it is known that earthworm diversity is affected by this (González et al. 1996, Didden 2001, Curry 2004, Feijoo et al. 2011).

Data will be analysed using a linear mixed effects model framework using ‘lme4’ in R. The random effect structure will account for differences between different datasets, such as sampling methodology, researcher error/biases, as well as differences from study location. Based on suggestions in De Palma et al. 2018 and approaches used by Soliveres et al. 2016, models will first be constructed with measures of diversity (species richness, biomass and abundance) as response variables, with all environmental variables (soil properties, climate) and human impacts (land use/habitat cover) as predictors. The residuals from the first model will then be modelled as a function of time, thus checking, after accounting for environment and human impact, whether the residuals are dependent on time (Freckleton 2002). If the residuals are dependent on time, then biodiversity is changing over time more than can be explained by human impacts and other environmental variables. Additional analysis could include modelling time with all other
variables, to obtain parameter estimates, as well as focusing on how time-series length affects the results, as it is expected that time-series length will vary greatly, with most datasets sampling either twice over a couple of years or twice over many years.

We will also investigate alternative modelling approaches. One potential alternative will be the use of structural equation models (SEMs; Grace 2008, Eisenhauer et al. 2015, Shipley 2016), thereby investigating the direct effect of time on earthworm diversity as opposed to the direct and indirect effects of other variables (climate, soil properties and human impacts, for example, land use; Fig. 3). Model specification will depend on data availability, as SEMs will need to be fitted to each dataset and modelling implemented in the R package ‘piecewiseSEM’ (Lefcheck 2016) which is capable of handling complex SEM structures, including composite variables. As the direction of the pathways are not always clear (i.e. the relationship between climate and land use; Fig. 3), multiple structures will be tested.

![Figure 3](image)

Figure 3: Structural Equation “meta-“ model (Grace et al. 2012). We will investigate the exact structure of the SEM and variables. For example, composite variables or PCA analysis may be used for the ‘Climate’ and ‘Soil Properties’ as they are comprised of multiple variables. ‘Land use’, as a categorical variable, would likely be a composite variable. Analysis would compare the direct effect of time on biodiversity (red arrow) with other indirect pathways.

2.3.4 WP4 - Ecosystem functions

Objectives

Soil biodiversity is important for ecosystem functions and, consequently, changes in soil biodiversity may impact functions (e.g. Bradford et al. 2002, Bardgett and van der Putten...
2014). For example, reductions in decomposer diversity were associated with decreased decomposition and nutrient cycling (Heemsbergen et al. 2004, Handa et al. 2014). However, it is unclear how soil community composition changes with disturbances and consequently how any changes may impact ecosystem functions. Field and micro-/mesocosm experiments that examine changes in diversity and ecosystem function in response to disturbances often do not study the entire community. Therefore, the conclusions that can be drawn, in terms of how the soil responds to disturbance and how taxa contribute to ecosystem function, are limited to a subset of the community. The lack of primary literature means that we are unable to address this question across large scales or in a generalisable way.

In this WP, we will use specifically collected data on the soil community and ecosystem functions (from WP1) to investigate **how soil communities change with disturbance and how ecosystem functions change** as a result. We will investigate the following hypotheses (see Fig. 2): In a disturbed site, relative to an undisturbed site, all soil taxa groups are negatively impacted (decrease in community abundance) and ecosystem function is reduced (WP4-H1). Alternatively, not all soil taxa groups in a disturbed site will be negatively impacted (change in community composition) and ecosystem function is changed (WP4-H2). A third option is that in a disturbed site not all soil taxa groups are negatively impacted (change in community composition) but ecosystem function is maintained (WP4-H3).

**Methods**

Data will be collated on soil fauna and soil functions through the establishment of the global monitoring network in WP1. These data, collected using a standardised sampling protocol from sites distributed across the globe, will be analysed as part of this WP. A literature search will be performed to obtain recent published studies that meet all criteria. In addition, we are also aware of one large dataset (Biodiversity Exploratories) that has collected suitable data from across 150 sites in Germany. We will ask for access to this dataset to add to the analysis.

Suitable datasets will include data from two or more sites, which vary in their environmental conditions (e.g. land use, habitat cover, soil properties, climate). Each dataset will contain the sampled diversity (biomass and abundance) of soil macrofauna at the order-level (e.g. earthworms, beetles, centipedes/millipedes), as well as the biomass/abundance/diversity of the earthworm community. In addition, soil functions (decomposition, soil pH, aggregate stability, microbial C, microbial respiration, metabolic quotient (qO2)) at each site will also have been measured (discussed in WP1). As the coordinates of each site will be known, external data layers (such as SoilGrids and CHELSA climate data) can also be included in the analysis.

Although all the data will be standardised across the datasets, mixed effect models will be tested in case differences between datasets that may have arisen due to differences amongst multiple data collectors need to be accounted for. Models will be constructed for each ecosystem function. Predictor variables will include environmental variables that were
collected in situ, such as land use and habitat cover, as well as those acquired from global layers that may explain additional variance within the data, such as altitude.

As little is known about how the diversity and composition of soil communities will change across a range of environmental conditions, such as different land uses, habitat covers and climates, we will first investigate this using the collated dataset. Diversity measures of the soil community, such as biomass, abundance and composition (discussed below), will be used as response variables in the models, with information on the environmental conditions at each site used as predictor variables.

To test whether changes in composition affect ecosystem functions, models will also contain the community composition of the soil fauna at the site as a predictor variable. The suitability of community composition metrics for the analysis will be investigated, to ensure that the abundance or biomass of each taxonomic order is clearly captured. For example, Simpson’s Evenness (Magurran 2004) would be a possibility, as it calculates the evenness in terms of abundance/biomass across all species or orders in a sample whilst removing biases created when samples have differing number of species. Sites from different regions will inevitably have different numbers of orders, so the diversity measures used will need to account for this.

The ecosystem functions that we are measuring may not be strongly impacted by all of the macrofauna groups sampled. Earthworms are known to impact some of the ecosystem functions measured, for example, decomposition, microbial activity, potentially more so than the other soil fauna groups that may be sampled. Therefore, models will also be created using earthworm diversity and the composition of the earthworm functional groups as the predictor variables (along with predictor variables mentioned previously) of the ecosystem functions. Further investigations into the effect of different soil fauna orders can then be done post hoc if needed.

**Deliverables**

**WP1**: At least one paper will be written to highlight and promote the creation of this monitoring network. In order to increase the participation in the under-represented countries, the paper will be submitted to a continent-specific, peer-reviewed journal. A more general paper, detailing the creation of the network and the subsequent database, will also be submitted on completion of the project to an international, peer-reviewed journal.

**WP2**: This synthesis analysis will allow us to write at least one paper in an international, peer-reviewed journal on the changes in earthworm communities as a response to human impacts.

**WP3**: This synthesis analysis will result in at least one paper in an international, peer-reviewed journal showing how earthworm communities are changing over time, adding to the existing debate on whether local diversity is changing over time.

**WP4**: This work package will result in at least one paper in an international, peer-reviewed journal. The paper will focus on using a collaborative network of researchers to address
how soil biodiversity is changing in response to disturbance and the effects on the ecosystem functions provided. All collaborating researchers will be invited to be an author on the paper.

**Data handling**

All data will be made publicly available with the respective publication using common data repositories, such as Dryad (http://datadryad.org/) and Pangaea (https://www.pangaea.de/), and assigned a DOI. Further, all data will be submitted to the Global Soil Biodiversity Database.

**Additional information**

Protocols for WP1 (See Supplementary Materials)

**Project requirements**

**Employment status information**

Phillips, Helen, Post-Doctoral Researcher, (currently based at the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany, (DFG FZT 118))

Eisenhauer, Nico, Full Professor (W3), permanent

**First-time proposal data**

Phillips, Helen

German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

**Composition of the project group**

*Dr. Simone Cesarz*, permanent post-doc in the Eisenhauer lab, specialist in soil chemical analyses. She will provide additional support and expertise for laboratory work. *Anja Zeuner*, technician within the Eisenhauer lab. She will conduct some of the laboratory work and provide assistance in the lab, if needed. *Svenja Haenzel*, Foreign Language Secretary in the Eisenhauer lab, will support the project by helping with contracting the student helper and with providing the necessary shipping documents and permits.
Cooperation with other researchers

Researchers with whom you have agreed to cooperate on this project

**Dr. Ika Djukic**, leads the global TeaComposition project (Djukic et al. 2018) and has expertise in soil microbial diversity and functions. **Dr. Carlos Guerra** currently leading an iDiv project collecting global soil samples from researchers in the TeaComposition project. In addition, expertise in future scenarios of global change and global biodiversity policy. **Dr. George Brown**, expertise in earthworm ecology and taxonomy, large database on tropical earthworms and important contact point in South America (Brazil). **Dr. Patrick Lavelle**, expertise in soil biodiversity and ecosystem function, leads global database on soil macrofauna with > 2,000 sampling locations (Orgiazzi et al. 2016, Cameron et al. 2018).

Researchers with whom you have collaborated scientifically within the past three years

Prof. Andy Purvis and all members of the PREDICTS project (http://www.predicts.org.uk/pages/team.html). Dr. Nicholas Fisichelli, Dr. Lee Frelich, Prof. Sarah Hobbie, Dr. Forest Isbell, Prof. Dr. Edward Johnson, Dr. Pete Manning, Prof. Peter Reich, Prof. Matthias Rillig, Prof. Jacques Roy, Prof. David Tilman, Prof. Wim van der Putten, Prof. Alexandra Wright, all PIs of iDiv (https://www.idiv.de/groups_and_people/members.html) and the Jena Experiment (http://www.the-jena-experiment.de/Members.html) and all editors of Pedobiologia – Journal of Soil Ecology.

Scientific equipment

A laboratory, with all necessary equipment for WP1, is available at iDiv (within the Experimental Interaction Ecology group led by Prof. Nico Eisenhauer). In addition, iDiv provides a high-performance computer cluster (HPC) and highly skilled IT support staff, that would be available should the analysis in WP2-4 require additional computational power. iDiv would provide the perfect infrastructure for the present project and no further equipment is requested.

References

- Araújo ASF, Borges CD, Tsai SM, Cesarz S, Eisenhauer N (2014) Soil bacterial diversity in degraded and restored lands of Northeast Brazil. Antonie van Leeuwenhoek,


Supplementary material

Suppl. material 1: SoilFaUNa Sampling Protocol [doι]

Authors: Helen RP Phillips
Data type: Sampling Protocol
Brief description: Sampling protocol to ensure standardised methods for researchers sampling soil macrofauna (and ecosystem functions) in a new site.
Download file (247.83 kb)