



## Drivers of earthworm incidence and abundance across European forests



Hans De Wandeler<sup>a</sup>, Rita Sousa-Silva<sup>a</sup>, Evy Ampoorter<sup>b</sup>, Helge Bruelheide<sup>c,d</sup>,  
 Monique Carnol<sup>e</sup>, Seid M. Dawud<sup>f</sup>, Gabriel Dănilă<sup>g</sup>, Leena Finer<sup>h</sup>,  
 Stephan Hättenschwiler<sup>i</sup>, Martin Hermy<sup>a</sup>, Bogdan Jaroszewicz<sup>j</sup>, François-Xavier Joly<sup>i</sup>,  
 Sandra Müller<sup>k</sup>, Martina Pollastrini<sup>l</sup>, Sophia Ratcliffe<sup>m</sup>, Karsten Raulund-Rasmussen<sup>f</sup>,  
 Federico Selvi<sup>l</sup>, Fernando Valladares<sup>n</sup>, Koenraad Van Meerbeek<sup>a</sup>, Kris Verheyen<sup>b</sup>,  
 Lars Vesterdal<sup>f</sup>, Bart Muys<sup>a,\*</sup>

<sup>a</sup> Department of Earth and Environmental Sciences, KU Leuven, University of Leuven, Celestijnenlaan 200E Box 2411, BE-3001 Leuven, Belgium

<sup>b</sup> Forest & Nature Lab, Department Forest and Water Management, Ghent University, Geraardsbergsesteenweg 267, B-9090 Gontrode, Belgium

<sup>c</sup> German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, D-04103 Leipzig, Germany

<sup>d</sup> Martin Luther University Halle-Wittenberg, Institute of Biology/Geobotany and Botanical Garden, Am Kirchtor 1, D-06108 Halle (Saale), Germany

<sup>e</sup> Laboratory of Plant and Microbial Ecology, University of Liège, Botany Bât. 22, Quartier Vallée 1, Chemin de la Vallée 4, 4000 Liège, Belgium

<sup>f</sup> Department of Geosciences and Natural Resource Management, University of Copenhagen, Rolighedsvej 23, DK-1958 Frederiksberg C, Denmark

<sup>g</sup> University Stefan cel Mare of Suceava, Forestry Faculty, 13 Universitatii Street, 720229 Suceava, Romania

<sup>h</sup> Natural Resources Institute Finland, Box 68, FI-80101 Joensuu, Finland

<sup>i</sup> Centre of Evolutionary and Functional Ecology (CEFE UMR 5175, CNRS – University of Montpellier – University Paul-Valéry Montpellier – EPHE), 1919 route de Mende, 34293 Montpellier Cedex 5, France

<sup>j</sup> Białowieża Geobotanical Station, Faculty of Biology, University of Warsaw, Sportowa 19, 17-230 Białowieża, Poland

<sup>k</sup> Department of Geobotany, Faculty of Biology, University of Freiburg, Schänzlestraße 1, 79104 Freiburg, Germany

<sup>l</sup> Department of Agri-Food and Environmental Sciences, University of Firenze, Piazzale delle Cascine 28, 50144, Firenze, Italy

<sup>m</sup> Department of Special Botany and Functional Biodiversity, University of Leipzig, 04103 Leipzig, Germany

<sup>n</sup> Museo Nacional de Ciencias Naturales, CSIC, Serrano 115 dpdo, E-28006 Madrid, Spain

### ARTICLE INFO

#### Article history:

Received 19 January 2016

Received in revised form

4 May 2016

Accepted 8 May 2016

#### Keywords:

Boosted regression trees (BRT)

Distribution

Climate

Soil fauna

Litter nutrients

Soil characteristics

### ABSTRACT

Earthworms have a significant influence on the structure, composition and functioning of forest ecosystems, but in spite of their role as ecosystem engineers, little is known on the factors controlling their distribution across European forests. Optimised sampling techniques, as well as more advanced statistical tools and geographical information systems have facilitated studies at the landscape scale. But these, and even larger-scale studies, are scarce due to data limitations, taxonomic inconsistencies and practical issues in linking existing databases. In this continental-scale field-based study we used boosted regression tree modelling to identify and evaluate the relative importance of environmental factors explaining earthworm incidence (presence/absence) and abundance (density and biomass) in European forests. To parameterise our models earthworms were sampled in six forest landscapes along a latitudinal gradient from the boreal north to the Mediterranean south in spring or autumn of 2012, together with several environmental variables. Earthworms were sampled using a combined method of mustard extraction and hand sorting of litter and a soil monolith, after which they were weighed and identified to functional group (epigeic, endogeic and anecic). We found that litter- and soil-related variables best explained earthworm incidence and biomass in European forests, leaving only a minor role to climate-related variables. Among the litter related variables, understory vegetation played an important role in explaining earthworm incidence and abundance. The relative importance of explanatory variables differed between models for incidence, density and biomass and between earthworm functional groups. Our results suggested that threshold values for soil C:N ratio, forest floor pH and understory plant biomass and plant nutrient concentrations have to be attained before earthworms can occur. Beyond these threshold values, variables like soil C:N ratio, tree litter C:P ratio and forest floor mass further

\* Corresponding author. Tel.: +32 16 32 97 26; fax: +32 16 32 97 60.

E-mail address: [Bart.Muys@kuleuven.be](mailto:Bart.Muys@kuleuven.be) (B. Muys).

explain earthworm biomass. Mechanisms behind these observations are discussed in the light of future earthworm distribution modelling at continental scale.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Earthworms are known as ecosystem engineers modifying the physical, chemical and biological soil properties (Blouin et al., 2013). They contribute to ecosystem functioning by changing soil porosity, by controlling the rate of organic matter decomposition and nutrient release, and consequently also indirectly influence primary production (Scheu and Wolters, 1991; Lavelle and Spain, 2001; Scheu, 2003; Edwards, 2004). Based on their ecology, Bouché (1977) classified earthworm species into three ecological groups, namely epigeic, endogeic and anecic earthworms. These functional groups have a different behaviour, and thus affect ecosystem functioning differently (Lavelle and Spain, 2001). Epigeic earthworms are rather small-sized species that live within the litter layer on the soil surface or within the uppermost part of the mineral soil and they feed on plant litter. Endogeic earthworms are geophagous species of generally intermediate size that live in a network of subhorizontal burrows in the mineral soil. Anecic earthworms finally, are large earthworm species that live in deep (semi) permanent vertical burrows in the mineral soil and feed on leaf litter that they drag into their burrows. Representing the greatest animal biomass component in most European forest soils (Lavelle and Spain, 2001), earthworm community composition and activity can have major consequences for the structure, composition and functioning of forest ecosystems (Hale et al., 2006; Lukac and Godbold, 2011). It is therefore highly relevant to know how environmental variables affect the incidence and abundance of earthworms (Schröder, 2008).

The composition of earthworm communities and their distribution can be studied at different spatial scales, e.g. local, landscape or continental scale. It is likely that factors influencing earthworm presence are scale-dependent (Schröder, 2008). At the local scale, earthworm incidence and abundance in forests has been shown to be influenced by soil moisture conditions (Whalen and Costa, 2003), soil texture (Fragoso and Lavelle, 1987), soil pH (Ma, 1984), litter quantity (Jordan et al., 2000; Nachtergale et al., 2002), and litter type (Peterson et al., 2001). At the landscape and continental scales, soil pH (Ammer et al., 2006; Moore et al., 2013) and litter type (Muys and Lust, 1992; Vahder and Irmeler, 2012) remain influential, but variables such as climate (Rutgers et al., 2015), land use history (Räty and Huhta, 2004) and dispersal possibilities (Räty and Huhta, 2004; Suárez et al., 2006; Shartell et al., 2013) become more important (Lavelle and Spain, 2001). Earthworm distribution at local scales has been studied in depth in most terrestrial ecosystems and bioclimatic regions worldwide (e.g. Fragoso and Lavelle, 1987; Peterson et al., 2001; Whalen and Costa, 2003; Suárez et al., 2006). More recently, a next generation of larger-scale studies has been stimulated by the availability of new sampling and species identification techniques, more advanced statistical tools and geographical information systems (Decaëns, 2010; Birkhofer et al., 2012; Shartell et al., 2013; Pansu et al., 2015). Despite these recent advancements and the importance for synthetic data analysis at larger spatial scales in soil ecology, only a few studies on earthworm distribution have exceeded the landscape scale (Lindahl et al., 2009; Decaëns, 2010; Rutgers et al., 2015). Studies at larger scales can provide new insights about the drivers of the abundance and composition of soil macrofauna

communities, but these are challenging for a number of reasons (Schröder, 2008; Cameron et al., in press). A standardised sampling executed over a large scale is costly, logistically, and often physically, demanding. Compiling datasets from different studies could therefore be an alternative solution. However, as Cameron et al. (in press) summarise, compiled datasets often suffer from data limitations due to differing sampling techniques, taxonomic inconsistencies and practical issues with linking and transferring databases.

In this paper we present the results of a continental-scale field-based study on earthworm incidence and abundance in mature forests. The objective of this research was to identify and evaluate the relative importance of a set of abiotic environmental variables (climate and soil parameters) and biotic variables such as vegetation composition and forest stand characteristics in explaining earthworm incidence and abundance in European forests. We evaluated the incidence and abundance of earthworm communities and the different functional groups of earthworms (epigeic, endogeic and anecic). As our sampling locations were spread along a latitudinal gradient from boreal to Mediterranean regions in Europe, covering distinct climate types (Peel et al., 2007), we hypothesised that earthworm community characteristics are primarily driven by climatic factors. Furthermore, since the three earthworm functional groups occupy different niches in the forest soil system and have distinct feeding behaviours (Bouché, 1977; Sheehan et al., 2008; Ferlian et al., 2014), we expected that the relative importance of the different control factors would differ between the three functional groups. We hypothesised that the incidence and abundance of endogeic species would primarily be explained by soil quality related factors, while that of epigeic and anecic species would be more influenced by litter quality parameters. This study extends on previous research at the landscape scale (Palm et al., 2013; Shartell et al., 2013; Marchán et al., 2015) by investigating earthworm community characteristics in mature forests at a continental scale. This research overcomes most of the aforementioned problems with large scale data sets, as all data were collected within a single study in the same standardised way, species were identified to the same taxonomic level and compiled into one single database.

## 2. Material and methods

### 2.1. Study sites

Sampling took place in the exploratory platform of the FunDiveEUROPE project (Baeten et al., 2013). This platform was designed to assess biodiversity-ecosystem function relationships along a tree species richness gradient in mature forests. The six studied forest landscapes (hereafter called sites) in this platform span most of the European bioclimatic gradient and represent major European forest types including the boreal forest (North Karelia, Finland), hemiboreal forest (Białowieża, Poland), temperate beech forest (Hainich, Germany), mountain beech forest (Râșca, Romania), thermophilous deciduous forest (Colline Metallifere, Italy) and Mediterranean-mixed forest (Alto Tajo, Spain) (Appendix 3, Fig. S1).

Each study site included between 28 and 43 selected plots (30 × 30 m) with different combinations of a fixed set of locally

dominant tree species. The established plots ranged in tree species diversity from one to five species per plot. Tree diversity however had no relevant effect on any of the earthworm variables and is therefore not further discussed. In total, the platform consisted of 209 plots with 16 target tree species, some of them occurring in multiple sites. The species pool comprised of conifers, deciduous broadleaves and evergreen broadleaves. For more details consult Baeten et al. (2013).

## 2.2. Field sampling

Earthworm sampling was carried out in spring 2012 for the sites in Italy, Germany and Finland, and in autumn 2012 for the sites in Poland, Romania and Spain. As advised in literature, we targeted our earthworm sampling in spring or autumn because of the humid soil conditions and positive temperatures (Berry and Jordan, 2001; Holmstrup, 2001; Baker and Whitby, 2003; Eggleton et al., 2009). In spite of this principle, our sampling campaign in Romania and Italy was characterised by extended drought periods prior to sampling, which influenced the sampling success. Plum and Filser (2005) estimated that it takes about half a year for an earthworm population to recover after a disturbance. Consequently, it could be that sites affected by recent droughts had not fully recovered yet, resulting in a lower earthworm abundance. Moreover, earthworms still in diapause (curled up 60–120 cm deep in soil (Sims and Gerard, 1999)) do not react to mustard suspension (Plum and Filser, 2005) and are out of reach for the 20 cm deep soil sampling. This could have resulted in an undersampling of anecic and endogeic earthworms.

Plots were divided in nine (10 × 10 m) subplots. In each plot, one earthworm sample was taken in the central subplot. Sampling close to tree stems was avoided and whenever possible performed in between different tree species. Earthworms were sampled by means of a combined method. First, litter was hand sorted over an area of 25 × 25 cm. Then, litter was removed over a larger area of 1 × 0.5 m in order to effectively apply an ethological extraction of earthworms using a mustard suspension (Valckx et al., 2011). Finally, hand sorting of a soil sample of 25 × 25 cm and 20 cm depth was performed in the middle of the 1 × 0.5 m area. Collected earthworms were preserved in ethanol (70%) for two weeks, transferred to a 5% formaldehyde solution for fixation (until constant weight), after which they were transferred back to ethanol (70%) for preservation of at least one month. Upon identification, all earthworms were individually weighed, including gut content, and classified into three functional groups, namely epigeic, endogeic and anecic species. Since preservation experiments showed that the weight of *Lumbricus rubellus* stabilises after two days of ethanol preservation (De Wandeler et al., unpublished results) we assumed the effect of a longer preservation period on weight was negligible. Results per unit area of the three sampling techniques were summed to determine the earthworm presence/absence, density and biomass per m<sup>2</sup>. For biomass recordings all earthworm biomass, even worm fragments, were considered. For density and presence/absence recordings, earthworm fragments were only considered if they had a head. By counting only heads, we prevented different fragments of the same earthworm being identified as multiple individuals.

For all 209 plots, a set of 35 explanatory variables were determined (Table 1). See supplementary material for details (Appendix 1). The explanatory variables used in this study described climate, soil, forest floor, understory, canopy and stand characteristics in the forest plots. Tree litter traits were calculated as community weighted means (CWM), both per plot (plot scale) and for a circular area with a 10 m radius surrounding the earthworm sampling location (local neighbourhood scale) and were represented by tree

litter C:N ratio, C:P ratio, specific leaf area index, cellulose, hemicellulose and polyphenol content. Plot scale litter trait CWMs are used in all reported models in the manuscript. In this study the specific leaf area index of tree litter (SLA) was found to be positively correlated with (i) litter nutritional quality variables (availability of base cations, readily available dissolved organic carbon and total dissolved nitrogen: Spearman's correlation coefficient,  $r_s$  resp. = 0.77, 0.61, 0.74;  $P < 0.001$ , respectively) and (ii) the regulation of forest floor microclimatic conditions (the capacity of litter to retain water:  $r_s = 0.70$ ;  $P < 0.001$ ). Consequently, SLA will be interpreted as such. The non-woody understory vegetation was determined and will be simply called understory vegetation in the remainder of the text.

## 2.3. Data analysis

We used Boosted Regression Trees (BRT) to identify the environmental variables most relevant in explaining the incidence and abundance of earthworms in mature European forests. BRT is a regression technique that can fit complex nonlinear relationships with threshold patterns, able to handle different types of explanatory variables, and it has a straightforward interpretable output (Elith et al., 2008; Aertsen et al., 2010). In addition, BRT has already proven to be a useful tool to model species distribution patterns (Leathwick et al., 2006; Elith et al., 2008; Cappelle et al., 2010; Palm et al., 2013).

Across all sampled forest plots, 24 different earthworm species were found. The majority of these species (67%) only occurred in one particular site and the most common species were found in 4 out of 6 sampled sites (Appendix 2, Table S1). Because of the high species turnover, we decided not to investigate species-specific responses, but instead we used the total earthworm incidence, density and biomass and the incidence, density and biomass of each functional group (i.e. epigeic, endogeic, anecic) as response variables in the BRT analyses. Explanatory variables are summarised in Table 1.

Data exploration indicated that the earthworm data were zero-inflated and had a strong positively skewed distribution. As a high proportion of zero observations in abundance models may impact model parameter estimates and consequently the response variable estimates (Lo et al., 1992), a two-step model approach called Delta approach or Hurdle model (Zuur et al., 2009) was used: first we modelled the probability of presence/absence with binomial BRTs (incidence models) and subsequently modelled the density and biomass using only non-zero observations with BRTs assuming a Gaussian error distribution (density and biomass models). In order to meet the requirements of homogeneity and normality of residuals, earthworm density and biomass were log<sub>10</sub> transformed prior to modelling.

To prevent our models from being affected by outliers or false negative values, some plots were removed from the analysis. Outliers and false negative values can arise from design and observer errors (Zuur et al., 2009). Design errors occurred in two plots where intense sheep grazing and defecation resulted in unusually high earthworm biomass. Observer errors occurred presumably due to inaccurate sampling practices in two plots or when the observer was unable to identify the sampled earthworm individual. This happened occasionally when small juveniles or damaged individuals could not be identified to species or functional group level. When the unidentified individuals contributed to 33% or more of total earthworm density or biomass, the plot was considered as an outlier with subsequent omission from the analysis, which was the case for only one plot in Italy. Since earthworm fragments were normally only considered as an existing individual if they had a head, false negative plots were created where only

**Table 1**  
Summary of the climate, soil and vegetation variables used in Boosted Regression Trees (BRT) analyses to explain the presence of earthworms in mature European forests. Mean (standard deviation), minimum and maximum for continuous variables; Median, minimum and maximum for rank variables; left blank in case of nominal variable.

Variable name	Description	Mean (SD)	Min.	Max.
<b>CLIMATE</b>				
Heat load index	Heat load index	0.82 (0.1)	0.52	1.02
Annual precipitation	Annual precipitation (mm)	641 (77)	484	819
Annual temperature	Annual mean temperature (°C)	7.7 (3.4)	1.3	14.0
Aridity	Aridity index	0.86 (0.21)	0.48	1.30
<b>SOIL</b>				
Slope	Slope of the plot (classes)	1	1	3
Bedrock	Bedrock weathering type (classes)			
Calcareous bedrock	Presence of calcareous bedrock (classes)			
Sand	Sand fraction (classes)	1	1	3
Silt	Silt fraction (classes)	2	1	3
Clay	Clay fraction (classes)	2	1	3
Soil drainage	Soil drainage condition (classes)	2	1	3
Soil Depth	Soil Depth (classes)	3	1	3
Soil type	acidic or buffered substrate (classes)			
Soil pH	pH of mineral soil at 10–20 cm (CaCl <sub>2</sub> )	4.62 (0.95)	3.39	7.31
Soil C:N	Soil C:N ratio at 10–20 cm of soil	16.2 (5.2)	6.8	35.6
Soil moisture	Soil volumetric water content (%)	31.64 (6.9)	17.10	46.87
<b>FOREST FLOOR</b>				
Humus type	Main humus types (classes)			
Forest floor pH	pH of ectorganic layer (CaCl <sub>2</sub> )	5.08 (0.82)	2.95	6.76
Forest floor mass	Dry mass of ectorganic layer (kg/m <sup>2</sup> )	2.102 (1.227)	0.404	9.851
Management	Current forest management (classes)			
Forest structure	Current forest structure (classes)			
LAI	Leaf area index (m <sup>2</sup> /m <sup>2</sup> )	4.38 (1.97)	0	9.10
Tree age distribution	Age distribution of tree canopy (classes)			
Understory biomass	Non-woody understory biomass (g/m <sup>2</sup> )	17.25 (20.54)	0	136.96
Understory C:N	C:N ratio of non-woody understory biomass	19.4 (7.6)	8.9	38.8
Understory C:P	C:P ratio of non-woody understory biomass	267.1 (164.1)	64.8	811.1
Tree richness	Number of tree species per plot	3 (1.1)	1	5
Tree evenness	Tree species evenness per plot	0.702 (0.275)	0	1
Litterfall biomass	Biomass of tree leaf litter (g/m <sup>2</sup> )	231.4 (112.95)	9.42	645.4
Litterfall C:N	CWM C:N of tree leaf litter	65.7 (35.8)	24.7	174.7
Litterfall C:P	CWM C:P of tree leaf litter	1131.0 (936.8)	195.9	4818.0
Litterfall cellulose	CWM cellulose of tree leaf litter (%)	18.52 (2.25)	12.61	26.26
Litterfall hemicellulose	CWM hemicellulose of tree leaf litter (%)	16.82 (1.48)	13.77	22.40
Litterfall polyphenols	CWM polyphenols of tree leaf litter (%)	39.9 (21.55)	9.7	109.5
Litterfall SLA	CWM SLA of tree leaf litter (cm <sup>2</sup> /g)	12.7 (6.1)	3.4	31.8

CWM = community weighted mean; SLA = specific leaf area.

unidentified earthworm fragments without heads were present. To minimise these false negatives and to maximise the number of plots used in the total incidence models, plots with only earthworm fragments were nevertheless recorded as containing earthworms. In our incidence models, up to 4 outliers and 4 false negative plots were recorded and removed. Up to 5 outliers were recorded and removed from the density and biomass models.

All analyses were performed in R (R Development Core Team 2014 version 3.1.1), using 'gbm' package version 2.1 (Ridgeway, 2006) and its extensions developed by Elith et al. (2008). BRT model parameters were set to constrain overfitting and optimise the number of regression trees for each model (mostly 2000–6000 trees), with a learning rate of 0.005–0.0005 and a bag fraction of 0.75–0.85. Tree complexity, i.e. the number of nodes in a tree, was set to 3 as recommended by Elith et al. (2008) for small datasets. Model fit and predictive performance after 10-fold cross validation of the biomass and density models were assessed by the percentage deviance explained ( $D^2$ ) (Littke et al., 2014), the correlation between observed and predicted values and the relative Root Mean Squared Error (rRMSE) (Aertsen et al., 2010). The incidence models were, in addition to the  $D^2$  and the correlation coefficient, assessed by the calculated area under the receiver-operating characteristic curve (AUC). The rRMSE was calculated by dividing the RMSE by the range of the response data [ $\text{RMSE}/(\text{Max}(\text{response})-\text{Min}(\text{response}))$ ] \*100.  $D^2$  is a measure for model fit (Buston and Elith, 2011), the RMSE is a measure of model accuracy (Aertsen et al., 2012) and the

AUC is a measure of model discrimination (Reineking and Schröder, 2006). Higher  $D^2$ , correlation and AUC scores, and lower rRMSE values indicate better performing models.

We minimised multicollinearity between explanatory variables in all models by the elimination of highly correlated variables that resulted in high variance inflation factors (i.e. individual VIF <0.5 and median VIF <0.3). In this selection process priority was given to retain variables that were, based on biological knowledge, assumed to be most directly linked to the response variable (Zuur et al., 2010) and to climate variables as in our continental-scale study, we wanted to test their importance as potential explanatory variable next to soil and forest floor variables. If both correlated variables were assumed to be equally closely related to the response variables, we dropped the explanatory variable with the lowest correlation with the response variable in order to maximise the explanatory power of the final models. As a consequence, the following explanatory variables were nearly always removed: tree leaf litter polyphenol content ( $r_s = -0.76$  with aridity;  $P < 0.001$ ), aromaticity of dissolved organic carbon ( $r_s = -0.78$  with tree leaf litter C:P;  $P < 0.001$ ), tree litter cellulose content ( $r_s = -0.7$  with SLA;  $P < 0.001$ ), tree litter C:N ratio ( $r_s = 0.67$  with tree litter C:P ratio;  $P < 0.001$ ), and understory C:N ratio ( $r_s = 0.72$  with understory C:P ratio;  $P < 0.001$ ). Collinearity between climate variables, namely between aridity and both annual mean temperature and annual precipitation, was handled differently to get a detailed understanding of the effects of different climate aspects. Final models

were run twice with two different sets of climate variables: 1) aridity, heat load index and annual mean temperature; and 2) annual precipitation and heat load index. As the main results of these models were similar we opted to only show the ‘aridity models’ and considered observed differences with the ‘annual mean temperature models’ in the discussion. In addition, the total incidence and biomass models were reanalysed with diversity and litter trait variables calculated at the local neighbourhood scale to test possible changes in relative importance of diversity and litter trait variables, but no consequential differences in relative importance were found (Appendix 3, Fig. S2).

In order to structure variable selection, two separate submodels with explanatory variables related to vertically distinct zones in the forest were developed. The first one contained forest floor, vegetation and stand characteristic variables and the second submodel contained all variables related with and situated below the ectorganic horizon, hereafter called: (i) forest floor submodel and (ii) soil submodel, respectively. These submodels were fitted with the *gbm.step* function and further simplified in a backwards stepwise manner by using the *gbm.simplify* function. Next we used a selection of explanatory variables from these simplified submodels whose relative importance exceeded 5%, together with the climate variables, to construct a final ecosystem model. In turn, final models were further simplified with the *gbm.simplify* function.

In the case of the density and biomass models for anecic earthworms we could not use the simplify function because of the small sample size (31 observations). However, we experienced from the other models that the relative importance of the most important variables hardly changed between different model runs, even after model simplification. Therefore the total anecic density and biomass models were built with the variables that had a relative importance of 5% or greater in the submodels. Since Kohavi (1995) reported that the cross-validation estimates of model predictive performance from small datasets can fluctuate, an average of five model runs was taken. From this model set, the model with the highest model performance estimates was retained to report model performance and partial dependence plots. We verified whether density models could be used as an indicator of earthworm biomass in case biomass values are not available. We realised that, despite the capabilities of BRT to limit overfitting, some could still occur in our partial dependence plots. To avoid evaluating potential spurious results we therefore only discussed the main trends in the partial dependence plots.

All reported final models had an acceptable level of collinearity between explanatory variables (i.e. Spearman correlations between variables <0.7, individual variance inflation factor (VIF) <0.5 and median VIF <0.3). All final models were checked for residual spatial autocorrelation by calculating Moran's I and spline correlograms. Moran's I was calculated using the ‘ape’ package (Paradis et al., 2015), after ‘Vincenty’ ellipsoid great circle distances between sampling points had been calculated with the ‘geosphere’ package (Hijmans et al., 2015). Spline correlograms with 95% pointwise bootstrap confidence intervals were calculated with the ‘ncf’ package (Bjornstad, 2013).

Lastly, we checked whether the results of the density models could be used as an indicator of earthworm biomass in the case that biomass values are not available in earthworm studies. The observed earthworm biomass was therefore plotted against predicted earthworm density.

### 3. Results

#### 3.1. Earthworm incidence and abundance

In the 209 sampled plots earthworm density ranged from 0 to

548 individuals per m<sup>2</sup>, with a biomass range from 0 to 204 g/m<sup>2</sup>. Average earthworm density and biomass were 90 earthworms/m<sup>2</sup> and 24 g/m<sup>2</sup>, respectively (Fig. 1). Earthworms were found in 81% of the plots. Endogeic species occurred in 57% of all sample plots, epigeic species in 50%, and anecic earthworms in only 16% of all sample plots (Fig. 2). Two functional groups were recorded in 33% of the plots and all three functional groups were recorded in 9% of the plots.

In mid-latitude forests (Germany and Poland) earthworms were found in all plots, in contrast with the northern and southern sites. In four of the six sampling sites endogeic earthworm species dominated (Fig. 2). However, in the forests of Finland and Poland, the epigeic species were more frequent than the other two functional groups. Anecic earthworms were generally less frequent, and they were completely absent in the forests of Spain.

#### 3.2. Relationship between environmental variables and earthworm incidence, biomass and density

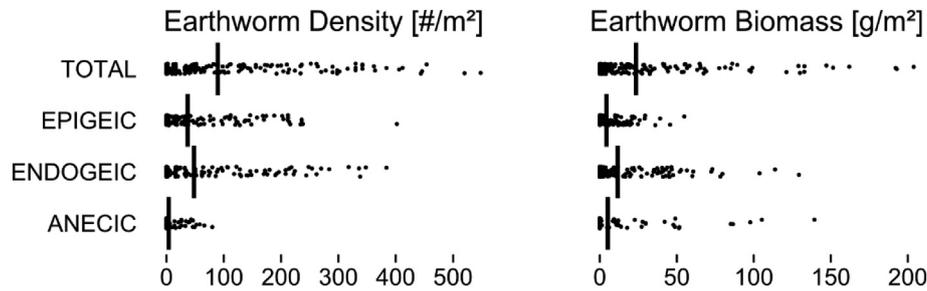
The BRT model fit statistics ranged from 0.89 to 0.74 for the correlation score and from 0.77 to 0.48 for the percentage deviance explained (D<sup>2</sup>). These statistics, together with the statistics after the 10-fold cross validation, are summarised in Table 2 and indicate confidence in the modelled results. Nevertheless, slight under- and overpredictions were observed for small and large biomass values, respectively (Appendix 3, Fig. S3). The weakest model statistics were found in the anecic biomass model. The final models did not show any significant residual spatial autocorrelation.

The total earthworm incidence was most strongly related to the C:N ratio and mass of the non-woody understory vegetation (35 and 32%, Fig. 3). Since C:N and C:P ratios of the understory vegetation were strongly positively correlated ( $r_s = 0.72$ ;  $P < 0.001$ ), earthworms appeared to prefer sites with an understory vegetation with a low C:N and C:P ratio (<20 and <300, respectively). Their presence was also most likely in forest floors with a relatively high pH (pH 5–6) and soils with a low C:N ratio (<20). When earthworms were present, their biomass was primarily related to soil C:N and tree litter C:P ratio, with higher earthworm biomass when the ratios were lower than 15 and 1000, respectively. Furthermore, a low forest floor mass and understory C:P ratio, and a high specific tree leaf area and aridity index, indicated a higher earthworm biomass.

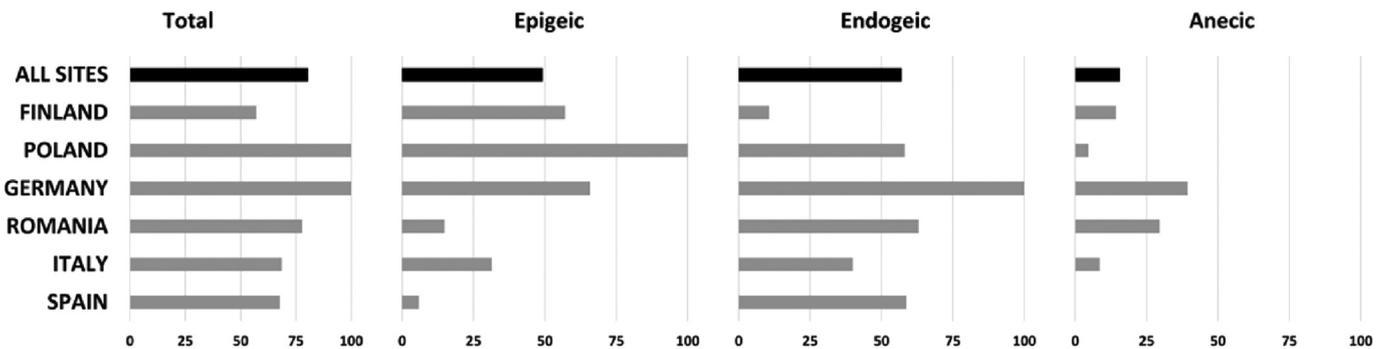
Both the incidence and biomass of epigeic earthworms were primarily related to the C:P ratio of the tree litter and non-woody understory vegetation, with a clearly higher probability of earthworm presence and also a higher earthworm biomass below C:P ratios of 200 and 1000, respectively (Fig. 4). Not only favourable litter quality, but also deep soils and sites with a low heat load index (N-NE facing slopes) promoted epigeic earthworm incidence.

In contrast, endogeic earthworm incidence was mainly related to humus type (29%), with a higher probability of earthworm presence in Amphi and Mull humus types (Fig. 5). Where endogeic earthworms occurred, their biomass was highest in plots with a low forest floor mass and a soil pH greater than 4. In addition, their incidence and biomass were higher at low soil C:N and both low litter C:N and C:P ratios. The epigeic and endogeic incidence models showed an opposite relation with soil C:N ratio around a common threshold of 12. The probability of the presence of epigeic earthworms was much higher in soils with a C:N ratio higher than 12, whereas the opposite was observed for the endogeic probability of presence (Fig. 6).

Anecic earthworm incidence was most related to the mass of the forest floor (63%), indicating higher probability of earthworm presence below 2 kg/m<sup>2</sup> dry matter and in more humid sites with an aridity index larger than 0.8 (Fig. 7). Anecics were most common



**Fig. 1.** Descriptive statistics of earthworm density (number/m<sup>2</sup>) and earthworm biomass (g/m<sup>2</sup>) across all individual sample plots. Each black dot indicates one of the 209 sampled plots distributed among six forest types across Europe. Black vertical bars indicate the respective mean values. To better visualise the data points, individual points are slightly vertically shifted (with the 'jitter' function in R) to prevent superposition.



**Fig. 2.** Earthworm incidence (%) in total (pooled across functional groups) (black), and for each functional group separately in the different forest sites (grey).

**Table 2**

Model performance indicators of the final boosted regression tree models, grouped by used response variable: area under the receiver-operating characteristic curve (AUC), percentage deviance explained (D<sup>2</sup>), correlation score (Corr) and relative Root Mean Squared Error (rRMSE) for model fit and after 10-fold cross-validation.

	Fit				Cross-validation			
	AUC	D <sup>2</sup>	Corr	rRMSE	AUC	D <sup>2</sup>	Corr	rRMSE
<b>Incidence models</b>								
Total	0.95	0.48	0.74		0.84	0.23	0.47	
Epigeic	0.97	0.60	0.84		0.89	0.40	0.68	
Endogeic	0.98	0.60	0.86		0.87	0.32	0.63	
Anecic	0.96	0.53	0.77		0.86	0.27	0.54	
<b>Biomass models</b>								
Total		0.69	0.85	12.18		0.47	0.70	16.01
Epigeic		0.77	0.89	10.38		0.52	0.74	14.92
Endogeic		0.65	0.83	12.01		0.34	0.58	16.65
Anecic		0.67	0.86	13.61		0.13	0.31	22.16

in sites with a soil pH range of 4.5–5.5, where the tree litter cellulose content was low (<17%) and the specific tree leaf area larger than 16 cm<sup>2</sup>/g.

Since annual mean temperature and annual precipitation were respectively strongly negatively and positively correlated with aridity, they were included in separate models. No remarkable differences between 'aridity models' and 'annual mean temperature models' were observed. However, the probability of total earthworm presence and total earthworm biomass were highest within a range of annual mean temperature of 6–9 °C (Appendix 3, Fig. S7).

Additionally, we identified and evaluated the environmental factors mostly related to earthworm density measures (total earthworm density and epigeic, endogeic and anecic earthworm density). When comparing these density model results with biomass model results, we observed fairly similar response curves

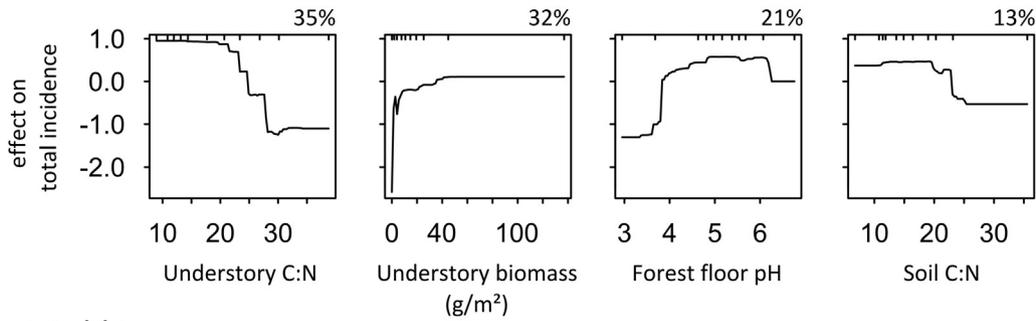
(Appendix 3, Fig. S4 and S5). Next to that, relatively strong positive relationships between predicted earthworm densities and observed earthworm biomass were observed (Appendix 3, Fig. S6). However in spite of these relatively high correlations, the relative importance (%) of the environmental variables between density and biomass models differed considerably.

## 4. Discussion

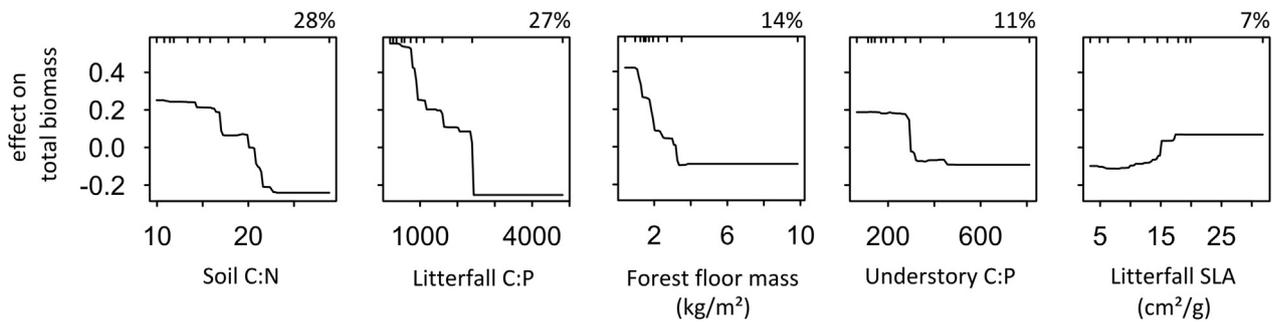
### 4.1. Environmental factors related to earthworm incidence and biomass in European forests

We identified the following explanatory variables as most important across the different models: forest floor mass, soil and forest floor pH, and soil and leaf litter nutrient concentrations. These results are consistent with studies at local (Hendriksen, 1990; Neiryck et al., 2000; Laganière et al., 2009) and landscape scales (Muys and Lust, 1992; Ammer et al., 2006; Shartell et al., 2013). The model results demonstrated that Boosted Regression Tree modelling is a useful technique to identify the relative importance of a set of different explanatory variables and how they influence earthworm incidence and abundance across different forests. Our models showed good explanatory power, though they could still be improved with additional sampling or eventually by merging our dataset with previously collected data. Doing so, it may become possible to create powerful distribution models to estimate predictive maps of total earthworm density, biomass and the three functional groups we distinguished across European forests. Our results may have been influenced by somewhat untypically low earthworm abundances in Italy and Romania resulting from unpredictable drought events during the respective sampling periods. Such stochastic effects of climate variability on the results could only be evaluated by repeated sampling campaigns over several years, which was beyond the scope of this study.

### Total incidence

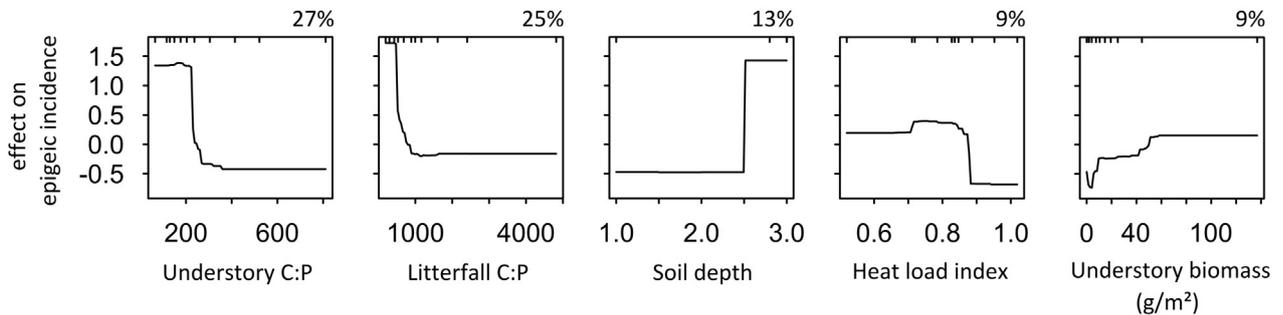


### Total biomass

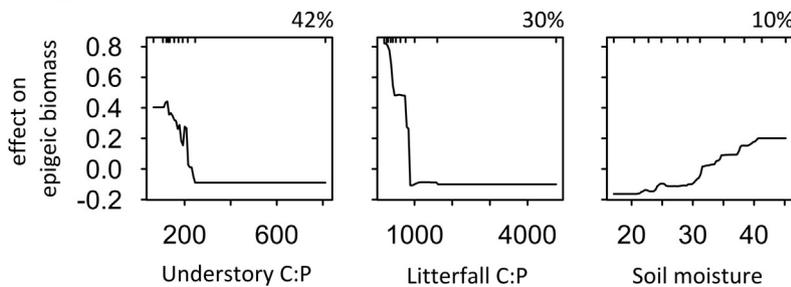


**Fig. 3.** Partial dependence plots of the five most influential variables that explain the total earthworm incidence (above) and total earthworm biomass (below) in European forests. The relative influence (%) of each explanatory variable is indicated above each graph. Each graph illustrates how the response variable (y-axis) is affected by a certain explanatory variable, after accounting for the average effects of all other variables in the model. The fitted functions are centred around the mean response value and plotted on a common scale (earthworm incidence on logit scale and the earthworm biomass on log scale). Hash marks along the top of the plots show the distribution of each explanatory variable in deciles.

### Epigeic incidence



### Epigeic biomass



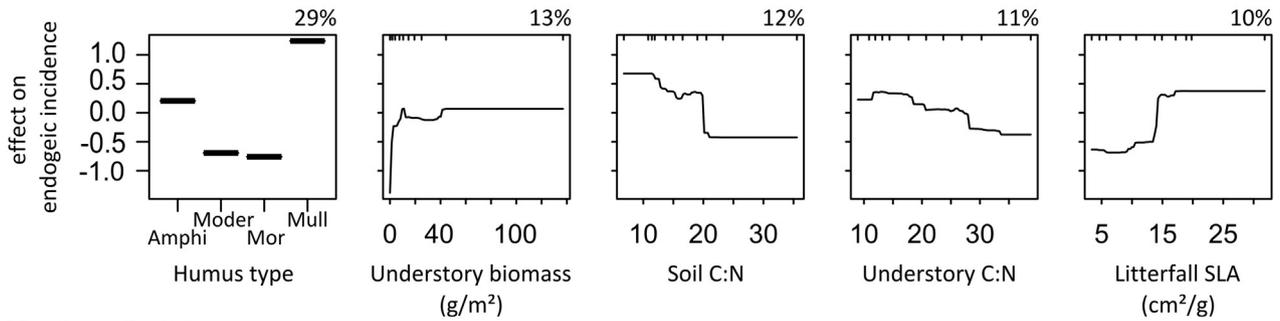
**Fig. 4.** Partial dependence plots of the five most influential variables that explain the epigeic earthworm incidence (above) and epigeic earthworm biomass (below) in European forests. The relative influence (%) of each explanatory variable is indicated above each graph. Each graph illustrates how the response variable (y-axis) is affected by a certain explanatory variable, after accounting for the average effects of all other variables in the model. The fitted functions are centred around the mean response value and plotted on a common scale (earthworm incidence on logit scale and the earthworm biomass on log scale). Hash marks along the top of the plots show the distribution of each explanatory variable in deciles.

#### 4.1.1. Climate

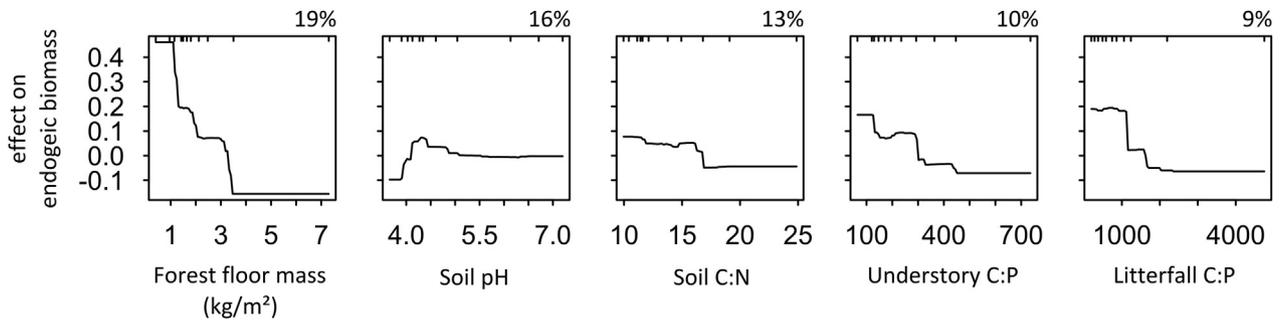
According to our model results, climatic factors play only a

minor role in explaining earthworm incidence and biomass across European forests. Apart from a higher probability of earthworm

## Endogeic incidence

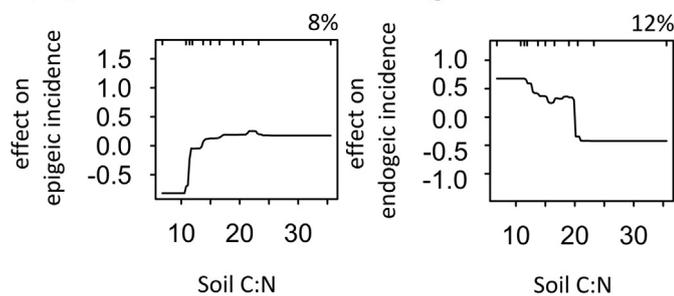


## Endogeic biomass

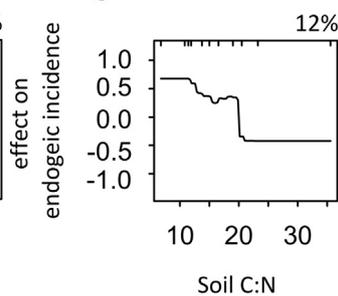


**Fig. 5.** Partial dependence plots of the five most influential variables that explain the endogeic earthworm incidence (above) and endogeic earthworm biomass (below) in European forests. The relative influence (%) of each explanatory variable is indicated above each graph. Each graph illustrates how the response variable (*y*-axis) is affected by a certain explanatory variable, after accounting for the average effects of all other variables in the model. The fitted functions are centred around the mean response value and plotted on a common scale (earthworm incidence on logit scale and the earthworm biomass on log scale). Hash marks along the top of the plots show the distribution of each explanatory variable in deciles.

## Epeigeic incidence



## Endogeic incidence



**Fig. 6.** Partial dependence plots of the epeigeic and endogeic earthworm presence probability for soil C:N ratio. The relative influence (%) of each explanatory variable is indicated above each graph. Each graph illustrates how the response variable (*y*-axis) is affected by a certain explanatory variable, after accounting for the average effects of all other variables in the model. Hash marks along the top of the plots show the distribution of each explanatory variable in deciles.

presence and higher earthworm biomass within a range of annual mean temperatures between 6 and 9 °C (Appendix 3, Fig. S7), there were only weak effects of climatic factors. In particular, the low predicted probability of the presence of anecic earthworms in arid conditions was unexpected, given the fact that many of these species show diapause behaviour as an adaptive strategy to drought (Lavelle, 1988). We may have missed the anecic earthworms in the drier sites, because large species occur at lower densities and because the sampling period may not have been optimal. The low probability of the presence of anecics in humid and cold conditions on the other hand confirms previous observations by Terhivuo (1989). The low probability of anecic

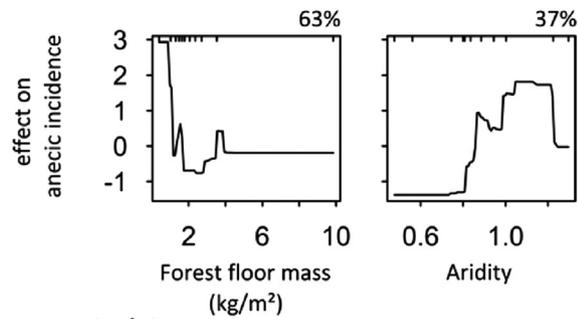
earthworm presence in forests of northern Europe is likely to be related to the unfavourable environmental conditions. Boreal forests are generally conifer-dominated forests on nutrient-poor podzolic soils (Barbati et al., 2007) with a low topsoil pH, which creates a toxic environment for anecic earthworms (Räty and Huhta, 2004).

## 4.1.2. Soil properties

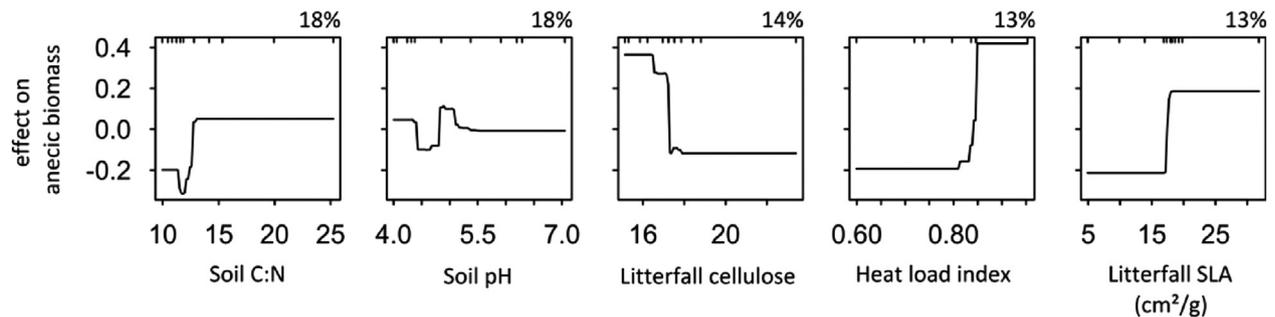
Soil C:N ratio was the most important soil property in our models related to total earthworm incidence as well as biomass. The entire earthworm community reached a biomass of 100–200 g/m<sup>2</sup> in fertile forest soils with soil C:N ratios less than 15. Such forest plots typically have Mull-like topsoils. The probability of earthworm presence decreased considerably above a C:N threshold value of 22, which corresponds to Mor-like topsoils. Similar observations were made by Muys and Lust (1992) in Flemish forests.

As could be assumed based on the feeding behaviour of endogeic earthworms (Ferlian et al., 2014), soil variables had a prominent role in predicting their distribution (Fig. 5), much more than in predicting the incidence and biomass of epeigeic and to some extent also the anecic earthworms. This corroborates our hypothesis that endogeic species would primarily be influenced by soil properties. Furthermore, the different ecology of the epeigeic and endogeic earthworm groups was emphasised by their opposite relationships with soil C:N ratio. High soil C:N ratios were attributed to poorly decomposable litter and consequent accumulation of incompletely humified organic material (Zanella et al., 2011) providing a habitat for epeigeic earthworms. Low soil C:N ratios, on the other hand, are related to Mull-like topsoils with high bioturbation activity of endogeic and anecic earthworms, leading to a rapidly disappearing litter layer, and thus, to fewer habitat opportunities and food for

## Anecic incidence



## Anecic biomass



**Fig. 7.** Partial dependence plots of the five most influential variables that explain the anecic earthworm incidence (above) and anecic earthworm biomass (below) in European forests. The relative influence (%) of each explanatory variable is indicated above each graph. Each graph illustrates how the response variable (y-axis) is affected by a certain explanatory variable, after accounting for the average effects of all other variables in the model. The fitted functions are centred around the mean response value and plotted on a common scale (earthworm incidence on logit scale and the earthworm biomass on log scale). Hash marks along the top of the plots show the distribution of each explanatory variable in deciles.

epigeic earthworms (Butt, 1998; Capowiez and Belzunces, 2001; Palm et al., 2013). It is striking that climatic variables had only weak effects in our models, but it also shows that, in relative terms, local humus type and the soil as a buffering system are of critical importance to earthworms (Edwards and Bohlen, 1996; Bonkowski et al., 1998).

In this context of predictive modelling it is important to understand that the choice of whether a variable is used as an explanatory or as a response variable is unrelated to ecological cause and effect relationships, and is rather driven by the objective to identify the factors that can explain earthworm presence and abundance in European forests. In this study for example, humus type is used as an explanatory variable for earthworm presence and abundance, but obviously earthworms themselves are a key driver in humus type formation.

### 4.1.3. Vegetation and soil-vegetation interactions

As hypothesised, leaf litter quality was an important factor in explaining earthworm incidence and biomass in European forests, i.e. tree leaf litter C:P, C:N and SLA (Hendriksen, 1990; Cortez, 1998; Holdsworth et al., 2012). Besides tree leaf litter, however, the quality of the non-woody understory vegetation (C:P and C:N ratio) was also important, but is commonly less considered. It is known, that low understory biomass C:N and C:P ratios support high soil microbial activity and growth (Bardgett, 2005) which in turn provides earthworms with more nutrients (Sampedro et al., 2006; Dungait et al., 2008; Ferlian et al., 2014). This suggests that understory vegetation may locally stimulate earthworm incidence, density and biomass via increased nutrient availability. This is supported by Kourtev et al. (1999) who reported that two exotic species in understory vegetation stimulated earthworm densities by increasing soil pH and nitrogen availability. It seems that non-

woody understory vegetation can significantly contribute to an increased food quality by decreasing the C:P ratio below a certain threshold (<200). It appears that epigeic earthworms can only occur below this threshold. Once this threshold is reached, epigeic earthworm biomass increases gradually with decreasing understory C:P ratio (Fig. 4). Similarly, we found a relatively high importance of tree leaf litter C:P ratio in the models for the total earthworm biomass and epigeic incidence and biomass. As phosphorus is generally present in low concentrations, it may limit microbial decomposer abundance and consequently the recycling of other important nutrients (Gosz et al., 1973; Tessier and Raynal, 2003; Blanes et al., 2012), which in turn could limit the nutrient availability for other organisms like earthworms.

While less important than leaf litter C:nutrient ratios, the SLA of the tree litter contributed significantly to 60% of our models with a relative importance ranging from 6 to 16% (Appendix 2, Table S2). In models where SLA had an impact, we identified a threshold value of 15 cm<sup>2</sup>/g (10 cm<sup>2</sup>/g for epigeics), with a higher probability of earthworm presence and higher biomass values above this threshold value. High SLA is typically positively correlated with high leaf nutrient concentrations (Poorter and De Jong, 1999; Hoffmann et al., 2005), which we also observed in our study. Accordingly, high nutrient concentrations associated with high SLA values may be the ultimate driver stimulating earthworm incidence and abundance (Muys and Lust, 1992; Reich et al., 2005; De Schrijver et al., 2012).

Forest floor mass was an important explanatory factor in the total earthworm biomass model (Fig. 3). The negative relationship between forest floor mass and earthworm biomass illustrates that high earthworm abundance typically leads to a rapid disappearance of the forest floor (Cortez and Bouché, 1998). Furthermore, forest floor mass correlated well with total earthworm biomass, but

not with earthworm incidence. Therefore forest floor mass does not allow us to predict earthworm incidence, however where earthworms are present, the forest floor mass is negatively correlated to earthworm biomass due to earthworm decomposition and bioturbation activity (Bohlen et al., 2004; Drouin et al., 2016).

It would have been interesting to also consider plant root related parameters, such as root biomass, root exudates and root nutrient concentrations (C:N, C:P). In forest ecosystems root biomass typically makes up a significant fraction of the total litter production (Vogt et al., 1986; Saugier et al., 2001; Leppälammil-Kujansuu et al., 2014), and seems to play an often overlooked, but important role (Pollierer et al., 2007). Even though root and foliar N and P are often correlated across species on a global scale, their relation within sites is not always clear (Hobbie, 2015). Further exploration of their importance, in addition to the aboveground predictors, is therefore important.

#### 4.2. Different factors explaining earthworm incidence and biomass

We observed that influential variables specifically explaining earthworm incidence often differed from factors explaining earthworm biomass. Similar observations were reported by Palm et al. (2013), where predictor contributions changed between incidence and density models. We believe that these differences are indicative of threshold values that must be reached before earthworms can occur (Curry, 2004).

The incidence of earthworms was generally high when the soil C:N ratio was lower than 23, where sufficient non-woody understorey vegetation was present with a low C:N ratio (<25), and also when the forest floor pH was higher than 4. Once these prerequisites for earthworm incidence were met, earthworm biomass could be more precisely explained by soil C:N, tree litter C:P and mass of the forest floor (Fig. 3). The epigeic functional group was atypical in this case, as the influential explanatory factors for incidence and biomass were the same. This may indicate that earthworm presence and biomass for epigeics are influenced by similar covariates whereas presence and biomass of the other groups were explained by different factors.

#### 4.3. Earthworm density as proxy for biomass

Our results indicate that earthworm density might be a good proxy for earthworm biomass in European forests. Nevertheless considerable differences in the relative importance of explanatory variables existed between density and biomass models. The largest differences between the density and biomass model results were observed between the endogeic and anecic groups (Figs. 5 and 7 and Fig. S4). This could be explained by the distortion in the linear relation between density and biomass for these functional groups. There is a greater range in earthworm biomass in plots with only few individuals (Appendix 3, Fig. S8). This range is enlarged by the presence of large K-selected endogeic and anecic earthworm species (Bouché, 1977; Dash, 2001) that occur in low densities primarily in Italy, but also in Spain and Romania. A direct confirmation of this inference can be observed in the endogeic and anecic density models that described a decreasing number of earthworms in more arid environments. Incorporation of either density or preferentially biomass models along with incidence models will provide a better understanding of earthworm distributions (Meynard and Quinn, 2007).

#### Acknowledgements

This research has received funding from the European Union's Seventh Programme (FP7/2007–2013) under grant agreement no.

265171. We thank Filippo Bussotti, Stefan Dondeyne, Michael Scherer-Lorenzen, Rudy Swennen and Olivier Bouriaud for advice and logistic support. We are grateful to Jorgen Opdebeeck, Eric Van Beek, Andrea Coppi, Fabrizio Fioravanti, Felix Berthold, Eric Van den Troost, Astrid Vannoppen, Jakub Zaremba, Dawid Zielinski, Johan Coremans, Elsa Varela, Ricardo Ontillera, Sam Ottoy, Paul Ilescu and Vasilie and his horse Stella for their technical assistance.

#### Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.soilbio.2016.05.003>.

#### References

- Aertsens, W., Kint, V., De Vos, B., Deckers, J., Van Orshoven, J., Muys, B., 2012. Predicting forest site productivity in temperate lowland from forest floor, soil and litterfall characteristics using boosted regression trees. *Plant Soil* 354, 157–172.
- Aertsens, W., Kint, V., van Orshoven, J., Özkan, K., Muys, B., 2010. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecol. Model.* 221, 1119–1130.
- Ammer, S., Weber, K., Abs, C., Ammer, C., Prielzel, J., 2006. Factors influencing the distribution and abundance of earthworm communities in pure and converted Scots pine stands. *Appl. Soil Ecol.* 33, 10–21.
- Baeten, L., Verheyen, K., Wirth, C., Bruelheide, H., Bussotti, F., Finer, L., Jaroszewicz, B., Selvi, F., Valladares, F., Allan, E., Ampoorter, E., Auge, H., Avacariei, D., Barbaro, L., Barnoiaea, I., Bastias, C.C., Bauhus, J., Beinhoff, C., Benavides, R., Benneter, A., Berger, S., Berthold, F., Boberg, J., Bonal, D., Braggemann, W., Carnol, M., Castagnyrol, B., Charbonnier, Y., Checko, E., Coomess, D., Coppi, A., Dalmaris, E., Danila, G., Dawud, S.M., de Vries, W., De Wandeler, H., Deconchat, M., Domisch, T., Duduman, G., Fischer, M., Fotelli, M., Gessler, A., Gimeno, T.E., Granier, A., Grossiord, C., Guyot, V., Hantsch, L., Hattenschwiler, S., Hector, A., Hermy, M., Holland, V., Jactel, H., Joly, F.X., Jucker, T., Kolb, S., Koricheva, J., Lexer, M.J., Liebergesell, M., Milligan, H., Muller, S., Muys, B., Nguyen, D., Nichiforel, L., Pollastrini, M., Proulx, R., Rabasa, S., Radoglou, K., Ratcliffe, S., Raulund-Rasmussen, K., Seiferling, I., Stenlid, J., Vesterdal, L., von Wilpert, K., Zavala, M.A., Zielinski, D., Scherer-Lorenzen, M., 2013. A novel comparative research platform designed to determine the functional significance of tree species diversity in European forests. *Perspect. Plant Ecol. Evol. Syst.* 15, 281–291.
- Baker, G.H., Whitby, W.A., 2003. Soil pH preferences and the influences of soil type and temperature on the survival and growth of *Aporrectodea longa* (Lumbricidae). *Pedobiologia* 47, 745–753.
- Barbati, A., Corona, P., Marchetti, M., 2007. European Forest Types. Categories and Types for Sustainable Forest Management Reporting and Policy. European Environment Agency Technical Report no. 9/2006. Copenhagen.
- Bardgett, R.D., 2003. The Biology of Soil: a Community and Ecosystem Approach. Oxford University Press, New York.
- Berry, E.C., Jordan, D., 2001. Temperature and soil moisture content effects on the growth of *Lumbricus terrestris* (Oligochaeta: Lumbricidae) under laboratory conditions. *Soil Biol. Biochem.* 33, 133–136.
- Birkhofer, K., Schöning, I., Alt, F., Herold, N., Klärner, B., Maraun, M., Marhan, S., Oelmann, Y., Wubet, T., Yurkov, A., Begerow, D., Berner, D., Buscot, F., Daniel, R., Diekötter, T., Ehnes, R.B., Erdmann, G., Fischer, C., Foessel, B., Groh, J., Gutknecht, J., Kandeler, E., Lang, C., Lohaus, G., Meyer, A., Nacke, H., Näther, A., Overmann, J., Polle, A., Pollierer, M.M., Scheu, S., Schloter, M., Schulze, E.-D., Schulze, W., Weinert, J., Weisser, W.W., Wolters, V., Schruppf, M., 2012. General relationships between abiotic soil properties and soil biota across spatial scales and different land-use types. *PLoS One* 7, e43292.
- Bjornstad, O.N., 2013. Package 'ncf'. Version 1.1-5. <https://cran.r-project.org/web/packages/ncf/ncf.pdf>.
- Blanes, M.C., Emmett, B.A., Viñeola, B., Carreira, J.A., 2012. Alleviation of P limitation makes tree roots competitive for N against microbes in a N-saturated conifer forest: a test through P fertilization and <sup>15</sup>N labelling. *Soil Biol. Biochem.* 48, 51–59.
- Blouin, M., Hodson, M.E., Delgado, E.A., Baker, G., Brussaard, L., Butt, K.R., Dai, J., Dendooven, L., Peres, G., Tondoh, J.E., Cluzeau, D., Brun, J.J., 2013. A review of earthworm impact on soil function and ecosystem services. *Eur. J. Soil Sci.* 64, 161–182.
- Bohlen, P.J., Scheu, S., Hale, C.M., McLean, M.A., Migge, S., Groffman, P.M., Parkinson, D., 2004. Non-native invasive earthworms as agents of change in northern temperate forests. *Front. Ecol. Environ.* 2, 427–435.
- Bonkowski, M., Scheu, S., Schaefer, M., 1998. Interactions of earthworms (*Octolasion lacteum*), millipedes (*Glomeris marginata*) and plants (*Hordelymus europaeus*) in a beechwood on a basal hill: implications for litter decomposition and soil formation. *Appl. Soil Ecol.* 9, 161–166.
- Bouché, M.B., 1977. Stratégies lombriciennes. In: Lohm, U., Persson, T. (Eds.), *Soil Organisms as Component of Ecosystems*. Ecological Bulletins, Stockholm, pp. 122–132.
- Buston, P.M., Elith, J., 2011. Determinants of reproductive success in dominant pairs

- of clownfish: a boosted regression tree analysis. *J. Anim. Ecol.* 80, 528–538.
- Butt, K.R., 1998. Interactions between selected earthworm species: a preliminary laboratory-based study. *Appl. Soil Ecol.* 9, 75–79.
- Cameron, E.K., Decaëns, T., Lapiéd, E., Porco, D., Eisenhauer, N., 2015. Earthworm databases and ecological theory: synthesis of current initiatives and main research directions. *Appl. Soil Ecol.* <http://dx.doi.org/10.1016/j.apsoil.2015.11.012>. [http://ac.els-cdn.com/S0929139315301281/1-s2.0-S0929139315301281-main.pdf?\\_tid=f61e6354-1cf7-11e6-b1de-00000aab0f27&acdnat=1463576379\\_81b2ccb477da24a0316446cb714c8f09\\_inpress](http://ac.els-cdn.com/S0929139315301281/1-s2.0-S0929139315301281-main.pdf?_tid=f61e6354-1cf7-11e6-b1de-00000aab0f27&acdnat=1463576379_81b2ccb477da24a0316446cb714c8f09_inpress).
- Capowiez, Y., Belzunces, L., 2001. Dynamic study of the burrowing behaviour of *Aporrectodea nocturna* and *Allolobophora chlorotica*: interactions between earthworms and spatial avoidance of burrows. *Biol. Fertil. Soils* 33, 310–316.
- Cappelle, J., Girard, O., Fofana, B., Gaidet, N., Gilbert, M., 2010. Ecological modeling of the spatial distribution of wild waterbirds to identify the main areas where avian influenza viruses are circulating in the Inner Niger Delta, Mali. *EcoHealth* 7, 283–293.
- Cortez, J., 1998. Field decomposition of leaf litters: relationships between decomposition rates and soil moisture, soil temperature and earthworm activity. *Soil Biol. Biochem.* 30, 783–793.
- Cortez, J., Bouché, M.B., 1998. Field decomposition of leaf litters: earthworm-microorganism interactions—the ploughing-in effect. *Soil Biol. Biochem.* 30, 795–804.
- Curry, J.P., 2004. Factors affecting the abundance of earthworms in soils. In: Edwards, C.A. (Ed.), *Earthworm Ecology*, second ed. CRC Press, Boca Raton, Florida, pp. 91–113.
- Dash, M.C., 2001. *Fundamentals of Ecology*. Tata McGraw-Hill, New Delhi.
- De Schrijver, A., De Frenne, P., Staelens, J., Verstraeten, G., Muys, B., Vesterdal, L., Wuyts, K., Van Nevel, L., Schelfhout, S., De Neve, S., 2012. Tree species traits cause divergence in soil acidification during four decades of postagricultural forest development. *Glob. Change Biol.* 18, 1127–1140.
- Decaëns, T., 2010. Macroecological patterns in soil communities. *Glob. Ecol. Biogeogr.* 19, 287–302.
- Drouin, M., Bradley, R., Lapointe, L., 2016. Linkage between exotic earthworms, understory vegetation and soil properties in sugar maple forests. *For. Ecol. Manag.* 364, 113–121.
- Dungait, J.A.J., Briones, M.J.I., Bol, R., Evershed, R.P., 2008. Enhancing the understanding of earthworm feeding behaviour via the use of fatty acid  $\delta^{13}C$  values determined by gas chromatography-combustion-isotope ratio mass spectrometry. *Rapid Commun. Mass Spectrom.* 22, 1643–1652.
- Edwards, C.A., 2004. *Earthworm Ecology*, second ed. CRC Press, Boca Raton, Florida.
- Edwards, C.A., Bohlen, P.J., 1996. *Biology and Ecology of Earthworms*. Chapman and Hall, London.
- Eggleton, P., Inward, K., Smith, J., Jones, D.T., Sherlock, E., 2009. A six year study of earthworm (*Lumbricidae*) populations in pasture woodland in southern England shows their responses to soil temperature and soil moisture. *Soil Biol. Biochem.* 41, 1857–1865.
- Eliith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802–813.
- Ferlian, O., Cesarz, S., Marhan, S., Scheu, S., 2014. Carbon food resources of earthworms of different ecological groups as indicated by  $^{13}C$  compound-specific stable isotope analysis. *Soil Biol. Biochem.* 77, 22–30.
- Fragoso, C., Lavelle, P., 1987. The earthworm community of a Mexican tropical rain forest (Chajul, Chiapas). In: Bonvicini, A.M., Omodeo, P. (Eds.), *On Earthworms*. Mucchi Modena, pp. 281–295.
- Gosz, J.R., Likens, G.E., Bormann, F.H., 1973. Nutrient release from decomposing leaf and branch litter in the Hubbard Brook Forest, New Hampshire. *Ecol. Monogr.* 43, 173–191.
- Hale, C.M., Frelich, L.E., Reich, P.B., 2006. Changes in hardwood forest understory plant communities in response to European earthworm invasions. *Ecology* 87, 1637–1649.
- Hendriksen, N.B., 1990. Leaf litter selection by detritivore and geophagous earthworms. *Biol. Fertil. Soils* 10, 17–21.
- Hijmans, R.J., Williams, E., Vennes, C., Hijmans, M.R.J., 2015. Package ‘geosphere’. Version 1.4-3. <https://cran.r-project.org/web/packages/geosphere/geosphere.pdf>.
- Hobbie, S.E., 2015. Plant species effects on nutrient cycling: revisiting litter feedbacks. *Trends Ecol. Evol.* 30, 357–363.
- Hoffmann, W.A., Franco, A.C., Moreira, M.Z., Haridasan, M., 2005. Specific leaf area explains differences in leaf traits between congeneric savanna and forest trees. *Funct. Ecol.* 19, 932–940.
- Holdsworth, A.R., Frelich, L.E., Reich, P.B., 2012. Leaf litter disappearance in earthworm-invaded northern hardwood forests: role of tree species and the chemistry and diversity of litter. *Ecosystems* 15, 913–926.
- Holmstrup, M., 2001. Sensitivity of life history parameters in the earthworm *Aporrectodea caliginosa* to small changes in soil water potential. *Soil Biol. Biochem.* 33, 1217–1223.
- Jordan, D., Hubbard, V.C., Ponder Jr., F., Berry, E.C., 2000. The influence of soil compaction and the removal of organic matter on two native earthworms and soil properties in an oak-hickory forest. *Biol. Fertil. Soils* 31, 323–328.
- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proc. 15th International Joint Conference on Artificial Intelligence*, 14, pp. 1137–1145.
- Kourtev, P.S., Huang, W.Z., Ehrenfeld, J.G., 1999. Differences in earthworm densities and nitrogen dynamics in soils under exotic and native plant species. *Biol. Invasions* 1, 237–245.
- Laganière, J., Paré, D., Bradley, R.L., 2009. Linking the abundance of aspen with soil faunal communities and rates of belowground processes within single stands of mixed aspen–black spruce. *Appl. Soil Ecol.* 41, 19–28.
- Lavelle, P., 1988. Earthworm activities and the soil system. *Biol. Fertil. Soils* 6, 237–251.
- Lavelle, P., Spain, A., 2001. *Soil Ecology*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Leathwick, J., Eliith, J., Francis, M., Hastie, T., Taylor, P., 2006. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Mar. Ecol. Prog. Ser.* 321, 267–281.
- Leppälampi-Kujansuu, J., Salemaa, M., Kleja, D.B., Linder, S., Helmsaari, H.-S., 2014. Fine root turnover and litter production of Norway spruce in a long-term temperature and nutrient manipulation experiment. *Plant Soil* 374, 73–88.
- Lindahl, A., Dubus, I.G., Jarvis, N., 2009. Site classification to predict the abundance of the deep-burrowing earthworm *L. Vadose Zone J.* 8, 911–915.
- Littke, K., Harrison, R., Zabowski, D., Ciol, M.A., Briggs, D.G., 2014. Prediction of Douglas-fir fertilizer response using biogeoclimatic properties in the coastal Pacific Northwest. *Can. J. For. Res.* 44, 1253–1264.
- Lo, N.C.-h., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on Delta-Lognormal Models. *Can. J. Fish. Aquat. Sci.* 49, 2515–2526.
- Lukac, M., Godbold, D.L., 2011. *Soil Ecology in Northern Forests: a Belowground View of a Changing World*. Cambridge University Press, New York.
- Ma, W.-c., 1984. Sublethal toxic effects of copper on growth, reproduction and litter breakdown activity in the earthworm *Lumbricus rubellus*, with observations on the influence of temperature and soil pH. *Environmental pollution series A. Ecol. Biol.* 33, 207–219.
- Marchán, D.F., Refoyo, P., Novo, M., Fernández, R., Trigo, D., Díaz Cosín, D.J., 2015. Predicting soil micro-variables and the distribution of an endogeic earthworm species through a model based on large-scale variables. *Soil Biol. Biochem.* 81, 124–127.
- Meynard, C.N., Quinn, J.F., 2007. Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *J. Biogeogr.* 34, 1455–1469.
- Moore, J.-D., Ouimet, R., Bohlen, P.J., 2013. Effects of liming on survival and reproduction of two potentially invasive earthworm species in a northern forest podzol. *Soil Biol. Biochem.* 64, 174–180.
- Muys, B., Lust, N., 1992. Inventory of the earthworm communities and the state of litter decomposition in the forests of Flanders, Belgium, and its implications for forest management. *Soil Biol. Biochem.* 24, 1677–1681.
- Nachtergale, L., Ghekiere, K., De Schrijver, A., Muys, B., Luysaert, S., Lust, N., 2002. Earthworm biomass and species diversity in windthrow sites of a temperate lowland forest. *Pedobiologia* 46, 440–451.
- Neiryneck, J., Mirtcheva, S., Sioen, G., Lust, N., 2000. Impact of *Tilia platyphyllos* Scop., *Fraxinus excelsior* L., *Acer pseudoplatanus* L., *Quercus robur* L. and *Fagus sylvatica* L. on earthworm biomass and physico-chemical properties of a loamy topsoil. *For. Ecol. Manag.* 133, 275–286.
- Palm, J., van Schaik, N.L.M.B., Schröder, B., 2013. Modelling distribution patterns of anecic, epigeic and endogeic earthworms at catchment-scale in agro-ecosystems. *Pedobiologia* 56, 23–31.
- Pansu, J., De Danieli, S., Puissant, J., Gonzalez, J.-M., Gielly, L., Cordonnier, T., Zinger, L., Brun, J.-J., Choler, P., Taberlet, P., Cécillon, L., 2015. Landscape-scale distribution patterns of earthworms inferred from soil DNA. *Soil Biol. Biochem.* 83, 100–105.
- Paradis, E., Bolker, B., Claude, J., Cuong, H.S., Desper, R., Durand, B., Dutheil, J., Gascuel, O., Heibl, C., Lawson, D., Lefort, V., Legendre, P., Lemon, J., Noel, Y., Nylander, J., Oppen-Rhein, R., Popescu, A.-A., Schliep, K., Strimmer, K., de Vienne, D., 2015. Package ‘ape’. Version 3.3. <https://cran.r-project.org/web/packages/ape/ape.pdf>.
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* 11, 1633–1644.
- Peterson, A.C., Hendrix, P.F., Haydu, C., Graham, R.C., Quideau, S.A., 2001. Single-shrub influence on earthworms and soil macroarthropods in the southern California chaparral. *Pedobiologia* 45, 509–522.
- Plum, N.M., Filser, J., 2005. Floods and drought: response of earthworms and potworms (*Oligochaeta*: *Lumbricidae*, *Enchytraeidae*) to hydrological extremes in wet grassland. *Pedobiologia* 49, 443–453.
- Pollierer, M.M., Langel, R., Körner, C., Maraun, M., Scheu, S., 2007. The underestimated importance of belowground carbon input for forest soil animal food webs. *Ecol. Lett.* 10, 729–736.
- Poorter, H., De Jong, R., 1999. A comparison of specific leaf area, chemical composition and leaf construction costs of field plants from 15 habitats differing in productivity. *New Phytol.* 143, 163–176.
- Räty, M., Huhta, V., 2004. Earthworm communities in birch stands with different origin in Central Finland. *Pedobiologia* 48, 283–291.
- Reich, P.B., Oleksyn, J., Modrzyński, J., Mrozinski, P., Hobbie, S.E., Eissenstat, D.M., Chorover, J., Chadwick, O.A., Hale, C.M., Tjoelker, M.G., 2005. Linking litter calcium, earthworms and soil properties: a common garden test with 14 tree species. *Ecol. Lett.* 8, 811–818.
- Reineking, B., Schröder, B., 2006. Constrain to perform: regularization of habitat models. *Ecol. Model.* 193, 675–690.
- Ridgeway, G., 2006. *Generalized Boosted Models: a Guide to the gbm Package*. Rutgers, M., Orgiazzi, A., Gardi, C., Römbeck, J., Jänsch, S., Keith, A.M., Neilson, R., Boag, B., Schmidt, O., Murchie, A.K., Blackshaw, R.P., Pérès, G., Cluzeau, D.,

- Guernion, M., Briones, M.J.L., Rodeiro, J., Piñeiro, R., Cosín, D.J.D., Sousa, J.P., Suhadolc, M., Kos, I., Krogh, P.-H., Faber, J.H., Mulder, C., Bogte, J.J., Wijnen, H.J.v., Schouten, A.J., Zwart, D.d., 2015. Mapping earthworm communities in Europe. *Appl. Soil Ecol.* 97, 98–111.
- Sampedro, L., Jeannotte, R., Whalen, J.K., 2006. Trophic transfer of fatty acids from gut microbiota to the earthworm *Lumbricus terrestris* L. *Soil Biol. Biochem.* 38, 2188–2198.
- Saugier, B., Roy, J., Mooney, H.A., 2001. Estimations of global terrestrial productivity: converging toward a single number. In: Roy, J., Saugier, B., Mooney, H.A. (Eds.), *Terrestrial Global Productivity*. Academic Press, San Diego, pp. 543–557.
- Scheu, S., 2003. Effects of earthworms on plant growth: patterns and perspectives. *Pedobiologia* 47, 846–856.
- Scheu, S., Wolters, V., 1991. Influence of fragmentation and bioturbation on the decomposition of <sup>14</sup>C-labelled beech leaf litter. *Soil Biol. Biochem.* 23, 1029–1034.
- Schröder, B., 2008. Challenges of species distribution modeling belowground. *J. Plant Nutr. Soil Sci.* 171, 325–337.
- Shartell, L., Lilleskov, E., Storer, A., 2013. Predicting exotic earthworm distribution in the northern Great Lakes region. *Biol. Invasions* 15, 1665–1675.
- Sheehan, C., Kirwan, L., Connolly, J., Bolger, T., 2008. The effects of earthworm functional diversity on microbial biomass and the microbial community level physiological profile of soils. *Eur. J. Soil Biol.* 44, 65–70.
- Sims, R.W., Gerard, B.M., 1999. Earthworms: notes for the identification of British species. In: Barnes, R.S.K., Crothers, J.H. (Eds.), *Synopses of the British Fauna (New Series)* Field Studies Council, Shrewsbury (no. 31 revised).
- Suárez, E., Tierney, G., Fahey, T., Fahey, R., 2006. Exploring patterns of exotic earthworm distribution in a temperate hardwood forest in South-Central New York, USA. *Landsc. Ecol.* 21, 297–306.
- Terhivuo, J., 1989. The Lumbricidae (Oligochaeta) of southern Finland: species assemblages, numbers, biomass and respiration. *Ann. Zool. Fenn.* 26, 1–23.
- Tessier, J.T., Raynal, D.J., 2003. Use of nitrogen to phosphorus ratios in plant tissue as an indicator of nutrient limitation and nitrogen saturation. *J. Appl. Ecol.* 40, 523–534.
- Vahder, S., Irmiler, U., 2012. Effect of pure and multi-species beech (*Fagus sylvatica*) stands on soil characteristics and earthworms in two northern German forests. *Eur. J. Soil Biol.* 51, 45–50.
- Valckx, J., Govers, G., Hermy, M., Muys, B., 2011. Optimizing earthworm sampling in ecosystems. In: Karaca, A. (Ed.), *Biology of Earthworms*. Springer, Berlin, pp. 19–38.
- Vogt, K.A., Grier, C.C., Vogt, D.J., 1986. Production, turnover, and nutrient dynamics of above- and belowground detritus of world forests. In: MacFadyen, A., Ford, E.D. (Eds.), *Advances in Ecological Research*. Academic Press, London, pp. 303–377.
- Whalen, J.K., Costa, C., 2003. Linking spatio-temporal dynamics of earthworm populations to nutrient cycling in temperate agricultural and forest ecosystems. In: *The 7th International Symposium on Earthworm Ecology · Cardiff · Wales · 2002*. *Pedobiologia*, 47, pp. 801–806.
- Zanella, A., Jabiol, B., Ponge, J.F., Sartori, G., De Waal, R., Van Delft, B., Graefe, U., Cools, N., Katzensteiner, K., Hager, H., 2011. *European Humus Forms Reference Base* (accessed 20.12.15.). <http://hal.archives-ouvertes.fr/hal-00541496/>.
- Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.
- Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* 1, 3–14.