RESEARCH ARTICLE

Soil macroinvertebrate communities: A world-wide assessment

Patrick Lavelle[1] | **Jérôme Mathieu**[1] | **Alister Spain**[2] | **George Brown**[3] | **Carlos Fragoso**[4] | **Emmanuel Lapied**[5] | **Adriana De Aquino**[6] | **Isabelle Barois**[7] | **Edmundo Barrios**[8] | **Maria Eleusa Barros**[1] | **Jose Camilo Bedano**[9] | **Eric Blanchart**[10] | **Mark Caulfield**[11,12] | **Yamileth Chagueza**[13] | **Jun Dai**[14] | **Thibaud Decaëns**[15] | **Anahi Dominguez**[9] | **Yamileth Dominguez**[16] | **Alexander Feijoo**[17] | **Patricia Folgarait**[18] | **Steven J. Fonte**[11] | **Norma Gorosito**[18] | **Esperanza Huerta**[19,20] | **Juan Jose Jimenez**[21] | **Courtland Kelly**[11] | **Gladys Loranger**[22] | **Robelio Marchão**[23] | **Raphael Marichal**[24] | **Catarina Praxedes**[25] | **Leonardo Rodriguez**[26] | **Guillaume Rousseau**[27] | **Laurent Rousseau**[1] | **Nuria Ruiz**[28] | **Catalina Sanabria**[29] | **Juan Carlos Suarez**[26] | **Jerôme Ebagnerin Tondoh**[30] | **Anne De Valença**[12] | **Steven J. Vanek**[11] | **Joel Vasquez**[31,32] | **Elena Velasquez**[13] | **Emily Webster**[33] | **Chi Zhang**[14]

[1]Institute of Ecology and Environmental Sciences IEES, Université Paris-Sorbonne Nouvelle, Paris, France

[2]Seaforth Road, Waihi Beach, New Zealand

[3]Forest Department, Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Colombo/PR, Brazil

[4]Red Biodiversidad y Sistemática, Instituto de Ecologia A.C., Xalapa, Mexico

[5]Taxonomia Biodiversity Fund, Paris

[6]Departamento de Agrobiologia, Empresa Brasileira de Pesquisa Agropecuária (Embrapa), Nova Friburgo, Brazil

[7]Red Ecología Funcional, Instituto de Ecologia A.C., Xalapa, Mexico

[8]Plant production and protection Division, FAO, Rome, Italy

[9]Geology/Natural Sciences Department, FCEFQyN, ICBIA-CONICET, Universidad Nacional de Rio Cuarto, Cordoba, Argentina

[10]Eco&Sols, Université de Montpellier, IRD, CIRAD, INRAe, Institut Agro, Montpellier, France

[11]Department of Soil and Crop Sciences, Colorado State University, Fort Collins, Colorado, USA

[12]Farming Systems Ecology, Wageningen University, Wageningen, The Netherlands

[13]Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia, Palmira, Colombia

[14]College of Natural Resources and Environment, South China Agricultural University, Guangzhou, China

[15]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 CNRS, Montpellier, France

[16]Facultad de Ciencias Básicas, Universidad del Atlántico, Barranquilla, Colombia

[17]Facultad de Ciencias Ambientales, Universidad Tecnológica de Pereira, Pereira, Colombia

[18]Unidad de Investigación en Interacciones Biológicas, Universidad Nacional de Quilmes, Buenos Aires, Argentina

[19]Departamento de Agroecologia, ECOSUR Campeche, Mexico

[20]Soil physics and land management group, Wageningen University and Research, Wageningen, The Netherlands

[21]Department of Biodiversity Conservation and Ecosystem Restoration, Pyrenean Institute of Ecology (ARAID-IPE-CSIC), Jaca, Spain

[22]Faculté des Sciences Exactes et Naturelles, Université des Antilles, Pointe á Pitre, Guadeloupe

[23]Centro de Pesquisa Agropecuária dos Cerrados (CPAC), Embrapa, Planaltina-DF, Brazil

[24]UMR ABSys, Centre International de Recherche en Agriculture et Développement, CIRAD, Montpellier, France

[25]Coordenação de zoologia, Museu Paraense Emilio Goeldi, Belém, Para, Brazil

[26]Facultad de Ingeniería, Universidad de la Amazonia, Florencia, Caquetá, Colombia

[27]Postgraduate Program in Agroecology, Maranhão State University, Sao Luis, Maranhão, Brazil

[28]Environments, Ecosystems and Biological Resources Department, French National Research Agency (ANR), Paris, France

[29]Departamento de Entomología, Universidad del Valle, UNIVALLE, Cali, Colombia

[30]UFR des Sciences de la Nature, Université Nagui Abrogoua, Abidjan, Côte d´Ivoire

[31]Dirección de Investigación en Diversidad Biológica Terrestre, Instituto de Investigaciones de la Amazonia Peruana, Iquitos, Peru

[32]Facultad de Ciencias Forestales, Universidad Nacional de la Amazonia Peruana, Iquitos, Peru

[33]Department of Plant Sciences, University of California, Davis, California, USA

Correspondence
Patrick Lavelle, Institute of Ecological and Environmental Sciences IEES, Université Paris-Sorbonne Nouvelle, Paris, France.
Email: plavelle48@gmail.com

Abstract

Aim: Macroinvertebrates comprise a highly diverse set of taxa with great potential as indicators of soil quality. Communities were sampled at 3,694 sites distributed worldwide. We aimed to analyse the patterns of abundance, composition and network characteristics and their relationships to latitude, mean annual temperature and rainfall, land cover, soil texture and agricultural practices.

Location: Sites are distributed in 41 countries, ranging from 55° S to 57° N latitude, from 0 to 4,000 m in elevation, with annual rainfall ranging from 500 to >3,000 mm and mean temperatures of 5–32°C.

Time period: 1980–2018.

Major taxa studied: All soil macroinvertebrates: Haplotaxida; Coleoptera; Formicidae; Arachnida; Chilopoda; Diplopoda; Diptera; Isoptera; Isopoda; Homoptera; Hemiptera; Gastropoda; Blattaria; Orthoptera; Lepidoptera; Dermaptera; and "others".

Methods: Standard ISO 23611-5 sampling protocol was applied at all sites. Data treatment used a set of multivariate analyses, principal components analysis (PCA) on macrofauna data transformed by Hellinger's method, multiple correspondence analysis for environmental data (latitude, elevation, temperature and average annual rainfall, type of vegetation cover) transformed into discrete classes, coinertia analysis to compare these two data sets, and bias-corrected and accelerated bootstrap tests to evaluate the part of the variance of the macrofauna data attributable to each of the environmental factors. Network analysis was performed. Each pairwise association of taxonomic units was tested against a null model considering local and regional scales, in order to avoid spurious correlations.

Results: Communities were separated into five clusters reflecting their densities and taxonomic richness. They were significantly influenced by climatic conditions, soil texture and vegetation cover. Abundance and diversity, highest in tropical forests (1,895 $\pm$ 234 individuals/m^2) and savannahs (1,796 $\pm$ 72 individuals/m^2), progressively decreased in tropical cropping systems (tree-associated crops, 1,358 $\pm$ 120 individuals/m^2; pastures, 1,178 $\pm$ 154 individuals/m^2; and annual crops, 867 $\pm$ 62 individuals/m^2), temperate grasslands (529 $\pm$ 60 individuals/m^2), forests (232 $\pm$ 20 individuals/m^2) and annual crops (231 $\pm$ 24 individuals/m^2) and temperate dry forests and shrubs (195 $\pm$ 11 individuals/m^2). Agricultural management decreased overall abundance by ≤54% in tropical areas and 64% in temperate areas. Connectivity varied with taxa, with dominant positive connections in litter transformers and negative connections with ecosystem engineers and Arachnida. Connectivity and modularity were higher in communities with low abundance and taxonomic richness.

Main conclusions: Soil macroinvertebrate communities respond to climatic, soil and land-cover conditions. All taxa, except termites, are found everywhere, and communities from the five clusters cover a wide range of geographical and environmental

conditions. Agricultural practices significantly decrease abundance, although the presence of tree components alleviates this effect.

KEYWORDS
communities, land cover, Macrofauna database, network analysis, soil ecology, soil macroinvertebrates

1 | INTRODUCTION

Soils are home to enormously biodiverse communities, estimated to make up as much as *c*. 40% of the total species on Earth (Orgiazzi et al., 2016). The greater part of this diversity remains unknown or undescribed, especially among the smaller organisms (Decaëns, 2010). Invertebrates are classified on the basis of size as micro-fauna (<0.2 mm), mesofauna (0.2–2 mm) and macrofauna (>2 mm) and are key determinants of self-organization in soils, a process whereby interactions at discrete scales among biotic and abiotic components create structures that have positive feedback on soil organism fitness (Lavelle et al., 2016; Perry, 1995). This important position is mediated through the control they exert on other soil community members and through their ecosystem engineering activities (Lavelle et al., 2016). We still lack a comprehensive study that characterizes their communities precisely at a world-wide scale. Macroinvertebrates, here defined as the invertebrates that can be seen with the naked eye, comprise a broad range of taxonomic groups with widely diverse ecologies and multiple influences on soil processes. Although their detailed roles as bio-chemical, physical and community engineers have been described in a wide range of publications (e.g., Blouin et al., 2013; Brussaard et al., 2006; Jouquet et al., 2011; Medina-Sauza et al., 2019), general models of soil function and management options still largely ignore these organisms (Barot et al., 2007; Bottinelli et al., 2015; Filser et al., 2016). This might be because of a general ignorance of their quantitative and functional roles in the soil. Such ignorance might be a consequence of their distribution across a wide range of publications and the distributions of data among many orders and families. It is even more surprising to see that soil biologists, specialists in other groups (microflora, microfauna or mesofauna), choose to ignore macrofauna in their attempts to propose ecological models of soil functioning (Guerra et al., 2021; Nielsen et al., 2015).

To overcome these limitations and propose this necessary comprehensive view of the importance and diversity of soil macroinvertebrate communities, a simple method for sampling macrofauna was developed to quantify the structure and diversity of their communities (ISO, 2011). Data collected over 30 years using this standard method have been accumulated world-wide and stored in a database. This paper provides a general overview of the macroinvertebrate communities studied and their major environmental determinants. The study is based on 3,694 sites distributed among 41 countries, collected in the Macrofauna database (Lavelle et al., 2022).

We first provide a general description of macroinvertebrate communities, the frequency of occurrence of the 16 main broad taxa, average and median densities world-wide. This should clearly inform on the abundance and diversity of this fauna at a global scale. We then searched for a general typology of the communities, looking for general patterns that could be taken into consideration in general conceptual and simulation models, and multivariate analyses were used for this purpose. Network analyses inside communities grouped according to the cluster analysis or land-cover groups completed these analyses.

Finally, we tested two general hypotheses. The first one states that local climatic conditions and natural and managed land cover are major drivers of the structure and population densities of soil faunal communities. A number of studies demonstrate such effects at local scales, but we do not know whether they are still operating at broader spatial scales and whether a hierarchy can be observed in the set of observed climate, soil and plant cover determinants (Lavelle et al., 1993). This would orientate practices aimed at enhancing this biological component and use macroinvertebrate communities as indicators of the different physical, chemical, morphological or biodiversity aspects of soil quality. The second hypothesis postulates that the general trends of reduced abundance, biodiversity and connectedness often observed in aboveground and belowground communities of microorganisms and soil micro- and mesoorganisms in agricultural versus non-agricultural soils (Morriën et al., 2017; Tsiafouli et al., 2015), although with different patterns depending on the taxa considered (Cameron et al., 2019), are paralleled in soil macroinvertebrate communities world-wide. Knowledge of these patterns will allow identification of the basic conditions that would conserve or enhance macroinvertebrate communities.

2 | MATERIALS AND METHODS

A standard method used at all sampling points provided a homogeneous set of data. Multivariate and network analyses provided the means to obtain a clear picture of the general patterns observed in macroinvertebrate communities and of their main drivers.

2.1 | Methodology: The ISO/TSBF protocol

In 1984, as part of the Decade of the Tropics UNESCO programme, the Tropical Soil Biology and Fertility (TSBF) programme was

launched, with the aim to "improve understanding of the role of biological and organic resources and their management for improved tropical soil fertility and the sustainability of land-use systems". Breaking with the taxonomic specialization that has led researchers to use taxon-specific methodologies, we proposed for this programme a macrofauna sampling method that involves hand sorting all the invertebrates recovered from 25 cm × 25 cm samples of the litter layer and the predominantly mineral soil, in 10 cm slices down to a depth of 30 cm (ISO, 2011; Velasquez, 2020). The definition of macrofauna was extended from the original ">2 mm size" to include all the invertebrates that could be seen with the naked eye.

This study is an attempt to synthesize the data collected in >50 independent projects in an opportunistic approach. Study sites were first concentrated in tropical areas where the TSBF programme operated, then extended progressively to other research sites and non-tropical areas as the method gained visibility and recognition. In the great majority of cases, the fieldwork was supervised by members of the team that initially proposed the method, or by people trained by them, which ensured a relative uniformity in the way of working and in the efficiency of the hand-sorting process.

In an evaluation of the efficiency of hand sorting conducted using large soil blocks, 1 m^2 × 30 cm, Lavelle et al. (1981) counted the organisms retained on sieves after previously hand-sorted soils were washed carefully through sieves and the residual fauna recovered. These authors and Jiménez et al. (2006) found that the efficiency of the hand-sorting technique varied greatly with organism size, colour and mobility and also differed among operators. Ruiz et al. (2008) and Velasquez (2020) found that it was important to follow the separation technique described in the ISO/TSBF method accurately and consistently to attain extraction efficiencies close to 100% for most invertebrates. Although differences in the individual efficiencies might, theoretically, affect the comparison among sites, we believe that they balance out owing to the high number of operators in all the >50 different projects, with an estimated 500 operators in total. Nonetheless, an effort has been made to provide people, whenever possible, with detailed technical documents and videos (Ruiz et al., 2008; Velasquez, 2020).

Nonetheless, the method is sensitive to the effects of differing soil conditions and land-use treatments on macroinvertebrate communities, has low operating costs and uses simple protocols. This has resulted in it becoming a standard method for the characterization of soil macroinvertebrate communities and a basis for the design and evaluation of synthetic indicators of soil biodiversity (Ruiz et al., 2011; Velasquez et al., 2007). It has been adopted as a tool for the description of soil quality in a large number of studies (Grimaldi et al., 2014; Lavelle et al., 2014; Rousseau et al., 2013; Velasquez and Lavelle, 2019). Although it was designed initially as part of a programme focused on tropical areas, it has since been adopted in large projects developed in Europe, China and the USA by local universities and research institutes.

At a given site, it is recommended that five samples 25 cm × 25 cm and 30 cm deep be taken, subdivided into four depth-based subsamples: the litter layer, 0–10, 10–20 and 20–30 cm. Ruiz et al. (2011)

showed that a minimum of three samples, separated laterally by 30 m to avoid autocorrelation, was enough to detect significant differences among a large set of natural and cultivated French soils. Subsamples from each separate layer of the litter and soil are kept in plastic bags and hand sorted immediately after the subdivision process described above. Invertebrates are preserved in alcohol.

Macroinvertebrates are generally identified to the ordinal level using simple keys.

2.2 | The Macrofauna database

All the data collected from sites sampled using the ISO/TSBF method have been stored in the Macrofauna database in an opportunistic approach. The ISO/TSBF method has been applied independently in a large number of different projects with different purposes since the earliest publication of the method, in 1989, and their data were also stored in the database. As a result of these dynamics, the data set has not been generated in an organized way, as designed from the beginning.

In 1999, however, the Macrofauna project developed within the framework of the International Biodiversity Year programme (IBOY; www2.nrel.colostate.edu/projects/iboy/whatandwhere.html), which brought together c. 40 people who had used the TSBF method, in an effort to optimize the method and share data within the framework of the Macrofauna database. A very detailed format was proposed, and the available data were stored. The data set used in the present study has been derived from the general database created at this time.

Sites were all sampled only once, and comparisons among types of plant cover were made according to a space-for-time sampling strategy. Access to the database was provided on the DRYAD platform (Lavelle et al., 2022). For each "site", data from one to five or more replicated blocks were pooled and densities expressed as individuals per square metre. Where the purpose of the study was to compare soils submitted to different types of land use in the same place, it was not considered advantageous to measure biomasses because they usually exhibit greater variances than equivalent population density data (Lavelle et al., 1981). Data sites were unevenly distributed globally, with substantial concentrations of sites sampled in Western Europe and tropical America, whereas large gaps were evident in Russia and North America (Figure 1). Latitude ranged from 55° S (Ushuaia, Argentina) to 57° N in Finland, elevation from 0 to 4,482 m in the Peruvian Andes, and mean annual temperatures from 5°C (Chuñuchuñuni in the Bolivian Andes) to >30°C in several sites of NE Brazil and Burkina Faso. Annual rainfall ranged from <1,000 mm in Queensland (Australia), Cordoba (Argentina) and Bolivia to >3,000 mm in Karnataka (India), Caqueta (Colombia), Sarawak (Indonesia) and Costa Rica.

For each site, the database recorded six geographical data: elevation, latitude, longitude, mean annual temperature and rainfall, and land cover. Latitude, elevation, temperature and rainfall were converted into between five and seven discrete classes, regularly
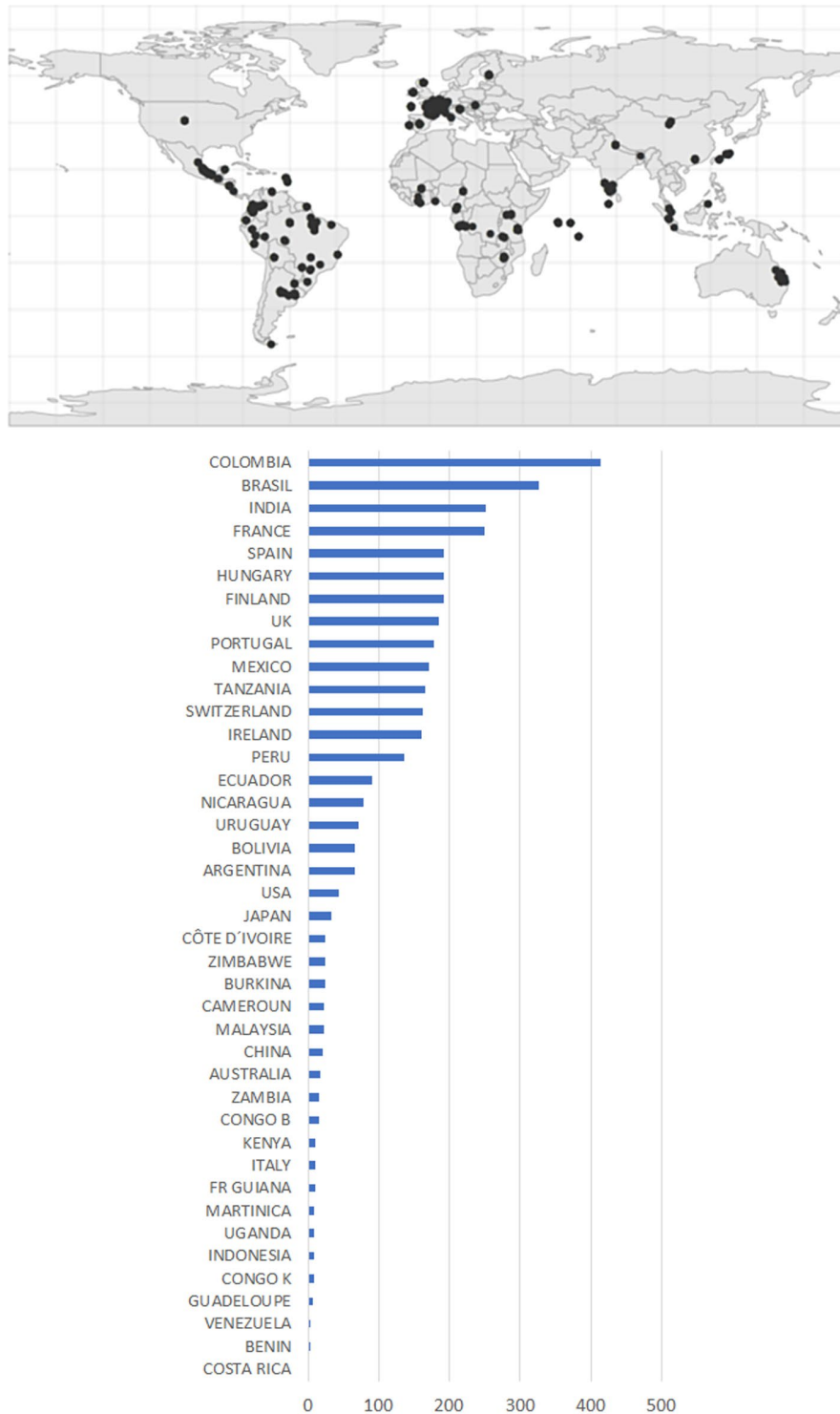
FIGURE 1 Top panel, distribution of the 3,694 sampling sites of the Macrofauna database; bottom panel, number of records per country

distributed according to the amplitude of their respective variations. The conversion of the data into discrete classes made it possible to homogenize data obtained from very diverse sources, those directly measured with great precision on site or those, less precise, drawn from cartographic documents. This approach is consistent with the global scale considered in the present study, which measures all variables at relatively general levels, such as orders for invertebrates or large latitudinal areas. It allows the use of robust statistical tools.

Eighteen variables recorded the properties of the macroinvertebrate community: taxonomic richness (the number of taxa found at the site); total density per square metre; and density per square metre of earthworms, ants, termites, Coleoptera, Chilopoda,

Diplopoda, Isopoda, Diptera, Blattaria, Hemiptera, Dermaptera, Lepidoptera, Orthoptera, Gastropoda, Arachnida and other groups. Soil texture information (provided in only 1,503 sites) was classified in three categories (sandy, silty or clayey), according to the dominant fraction in soil.

Sites were relatively more numerous within the latitudinal classes 45–60° and 0–15° and in the 500–1,000 mm range of annual rainfall (Figure S1).

Land cover, described in some detail in the original data sets, was reduced to 10 categories after preliminary analyses (not shown here). The first results obtained, some with a list of 21 different types of land cover inspired by the Food and Agriculture Organization of the United Nations and Corinne land-cover systems, showed a very large dispersion inside and among the groups identified in the factorial plans of the principal components analysis (PCA). This situation prevented the transmission of a clear and simple message at the global scale considered here. We propose a system that gives importance to environmental conditions that matter for soil macroinvertebrates: local climatic conditions (temperature and rainfall) and the importance of trees in the land cover. This will allow us to test the general hypothesis that the presence of trees in agricultural systems has positive effects on the diversity and abundance of macroinvertebrate communities.

This led us to distinguish natural systems [tropical rain forest (7.4% of the sites) and savannas (0.9%), temperate forests (deciduous, coniferous and mixed forests; 25.3% of the sites), grasslands (8.8%), and temperate dry forests and shrubs (2.1%)] and managed systems [temperate or tropical cropping systems, associated with trees (17.5%) or not (11.0 and 12.4%, respectively)]. A few sites that did not coincide with this classification, such as peatbogs, bare soil or urban soils, were classified into an "others" category.

2.3 | Data treatment

Data were checked before analysis, and errors of transcription were corrected. Histograms were plotted to examine the shapes of the frequency distributions of all taxa and total densities.

The measures of central tendency and variation are arithmetic means and standard errors, unless otherwise indicated.

Principal components analysis was used to ordinate the macroinvertebrate communities, described by its 18 respective variables, at the 3,694 sites and was performed with the *ade4* library for R (Dray & Dufour, 2007). We applied a correction factor according to the number of samples taken in each site (row weight, row.w, command in *ade4*).

A cluster analysis was then applied to the coordinates of the PCA performed on macroinvertebrate community data to separate them into a small set of distinct similar groups. This approach would generate a typology of the communities. We expected it to provide clues to identify the determinants of their composition and abundance. These were characterized further and linked to specific geographical

(latitude, elevation, mean annual rainfall and soil texture) and plant cover conditions.

Environmental data, latitude, elevation, plant cover and soil texture, were processed with a multiple correspondence analysis after expressing them in discrete classes. A global coinertia analysis between environmental and faunal data (Doledec & Chessel, 1994) allowed the significances of elevation, latitude, plant cover and soil texture effects on macroinvertebrate communities to be tested.

We then performed bias-corrected and accelerated (BCa) bootstrap tests (a particular case of redundancy analysis using a single explanatory variable) to evaluate separately the proportion of variance explained by each environmental variable and its respective significance. A bootstrapping procedure was associated with the BCa analysis to compensate for the unequal number of sites in the different classes. For each factor tested, we first removed strata with insufficient data, which were defined as levels with <100 observations. Then we resampled with replacement 1,000 times (number of bootstrap replicates) each stratum with the number of samples available in the smallest strata for the corresponding factor and computed the distribution of the bootstrapped variance explained by the factor for each bootstrap replicate. The detailed results are presented in the Supporting Information (Table S1).

Rainfall and elevation were both separated into seven classes from the lowest (0–500 mm/year and 0–500 m) to highest (>3,000 mm/year and 3,000 m) recorded values. Latitude was classified in five categories: 0–15, 15–30, 30–45, 45–60 and >60°, North or South. Eleven types of plant covers were identified after optimization of an initial set of 21 using preliminary statistical treatments.

Tests were made of the evidence of co-occurrence among taxonomic groups forming ecological networks in communities and their changes according to environmental conditions (Morueta-Holme et al., 2016). Each pairwise association of taxonomic units was tested against a null model considering local and regional scales in order to avoid spurious correlations. The null model was constructed by resampling the observed data separately within each relevant stratum (the "COVER8" and "CLUSTER" classification of sites according to their plant cover or grouping generated by the general cluster analysis performed on the dataset). We used a resampling scheme that preserves row and column sums. This procedure identifies significant positive and negative associations among taxa, which are used to define the local networks. The procedure was realized with the function "make netassoc network" from the *netassoc* (v.6.3) R package (Morueta-Holme et al., 2016). Clusters of taxonomic groups were identified using the *igraph* package (Csardi & Nepusz, 2006). Networks were then characterized by traditional metrics, such as connectance (total number of links divided by the total potential number of links), nestedness and modularity. Nestedness is a property of species assemblages where similar patterns of species occurrence are observed in low- and high-density communities (Atmar & Patterson, 1993). Modularity measures the division of a network into independent

subcommunities. This parameter is often assimilated with community resilience (Thebault & Fontaine, 2010).

3 | RESULTS

3.1 | General trends in macroinvertebrate communities

Macroinvertebrates occurred in 94% of the sites. Sites with no macroinvertebrate communities were found in semi-desert areas of Spain and Australia and in a few sites with intensive agricultural practices from Hungary. All taxonomic groupings except termites occurred at nearly all sites, although population densities differed strongly among taxa (Table 1). Earthworms were the most frequent taxon (71%), with Coleoptera (65%) and ants (53%).

3.2 | Frequency distributions of the population densities of the major taxa

The frequency distributions of the total population densities and of all the constituent taxa recorded were highly right skewed and platykurtic; consequently, large differences were apparent between the medians and the means for all taxa and for the total population densities (Table 1; Figure S2).

The three numerically dominant groups of organisms are the important soil ecosystem engineers: ants, termites and earthworms. Coleoptera, Myriapoda, Arachnida, Isopoda and Diptera larvae were

all less populous, all with mean densities <100 individuals/m². Lower values were found for all the remaining taxa and for the "others" category.

Although the mean density of the macroinvertebrate community was rather high, the median density for the whole data set was much lower than the mean and was zero for 13 of the 16 taxonomic groups. The data set is characterized by rather large proportions of zero values among the less populous taxa, right skews attributable to the highly aggregated distributions of many soil animal populations and generally high values of the excess kurtosis coefficient for all groups. This last feature might be attributable to underlying polymodal distributions.

Over the 3,694 sites sampled, macroinvertebrate communities contained a mean representation of 5.6 ± 0.07 of the taxonomic units presented in Table 1.

3.3 | A general typology of macroinvertebrate communities

Principal component analysis ranked the sites based on their overall densities and taxonomic richness along axis 1 (27.4% of variance explained; Figure 2a). The major ecosystem engineers (earthworms, ants and, to a lesser extent, termites) and two of the more important litter transformers (Myriapoda and Isopoda) were the most influential variables in defining this axis. Axis 2 (8.3% of variance explained) contrasted sites with high population densities of termites and Blattaria with sites possessing high densities of Coleoptera, Gastropoda, earthworms and Diptera larvae.

TABLE 1 Density and general statistical parameters of the main macroinvertebrate groups in the 3,694 sites represented in the Macrofauna database

| Taxonomic grouping	Frequency (%)	Arithmetic (mean ± SE)	Median	First quartile	Third quartile	Skewness coefficient	Kurtosis coefficient
Earthworms	70.7	123 ± 3.8	29.3	0.0	157.5	3.6	18.5
Ants	53.2	303 ± 29.3	6.9	0.0	99.0	17.5	397.9
Termites	26.4	187 ± 15.2	0.0	0.0	3.0	13.0	254.1
Coleoptera	65.1	50 ± 1.7	14.8	0.0	62.1	8.2	125.8
Chilopoda	35.8	13 ± 0.6	0.0	0.0	14.8	6.6	60.6
Diplopoda	33.1	21 ± 2.0	0.0	0.0	11.6	17.5	419.9
Isopoda	19.6	14 ± 1.7	0.0	0.0	0.0	16.3	359.6
Diptera	26.9	8 ± 0.5	0.0	0.0	2.2	11.9	214.3
Blattaria	13.6	2.3 ± 0.2	0.0	0.0	0.0	15.5	342.8
Hemiptera	19.1	3.2 ± 0.2	0.0	0.0	0.0	11.6	194.8
Dermaptera	9.0	1.3 ± 0.1	0.0	0.0	0.0	11.0	162.0
Lepidoptera	11.9	1.6 ± 1.6	0.0	0.0	0.0	12.9	256.1
Orthoptera	12.6	1.2 ± 1.2	0.0	0.0	0.0	9.5	147.0
Gastropoda	18.2	5.3 ± 0.4	0.0	0.0	0.0	14.8	387.6
Arachnida	48.8	17 ± 0.8	0.0	0.0	14.8	10.0	149.0
Others	15.2	12 ± 2.1	0.0	0.0	0.0	35.0	1,599.3
Total density	93.9	787 ± 36	250.2	78.4	630.0	11.8	206.7

(a)

(b)
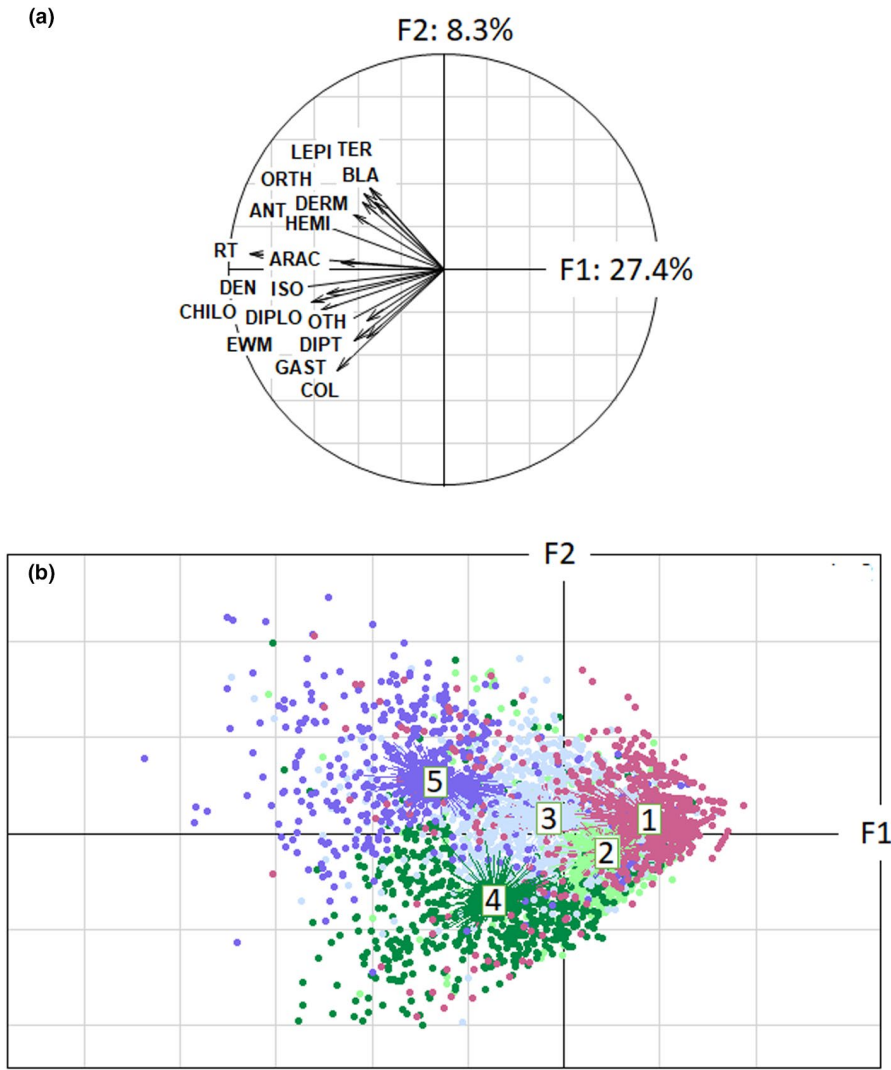
FIGURE 2 Projection of (a) macroinvertebrate group densities and community parameters and (b) sites grouped in five clusters in a plane defined by factors 1 and 2 (F1 and F2) of a principal components analysis of macroinvertebrate community data of the Macrofauna database. ANT, ants; ARAC, Arachnida; BLA, Blattaria; CHILO, Chilopoda; COL, Coleoptera; DEN, total density per square metre; DERM, Dermaptera; DIP, Diptera larvae; DIPLO, Diplopoda; EWM, earthworms; GAST, Gastropoda; HEMI, Homoptera Hemiptera; ISO, Isopoda; MYR, Myriapoda; ORTH, Orthoptera; OTH, other; RT, taxonomic richness (number of orders found in the sampling); TER, termites
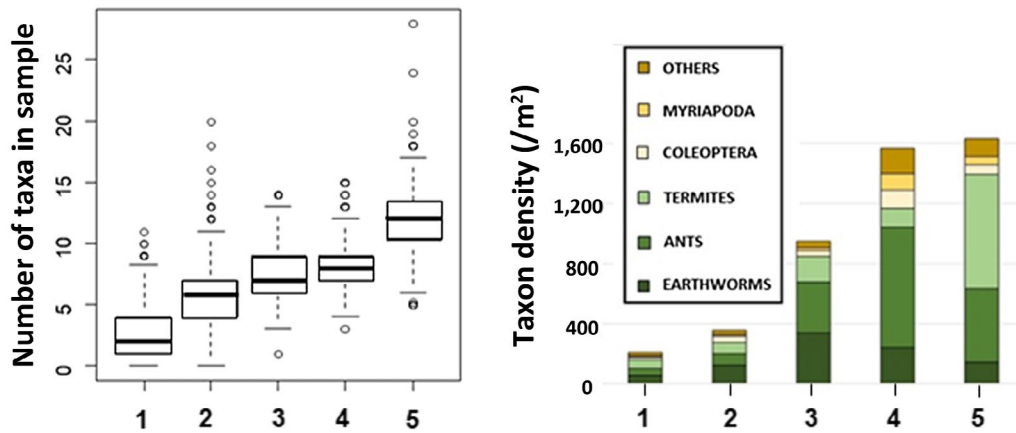

FIGURE 3 Left panel, taxonomic richness and right panel, composition of communities in the five clusters separated from the Macrofauna database

To explore the data set further, a cluster analysis was conducted on the set of macroinvertebrate data previously analysed with PCA (Figure 2b). The analysis separated five different clusters. Separation into these clusters explained 16.5% of total variance ($p < .001$; Table S1). There was a clear gradient in taxonomic richness and community density from clusters 1 to 5 (Figure 3). Although all clusters had sites from all the latitudinal classes, with the exception of the >60° class not represented in cluster 4, there was a clear trend for latitude to decrease from cluster 1 to 4. Cluster 5 exhibited a different trend, with rather evenly distributed latitude values (Figure S3).

All clusters comprised plant covers of all categories, although in different proportions, with relatively higher representation of temperate forests and grasslands in clusters 1–3 and tree-associated tropical crops and tropical forests in clusters 4 and 5 (Figure S4).

Soil macroinvertebrate communities were composed of similar organisms from cold temperate to equatorial soils, and throughout the five separated clusters (Figure 3). Termites were the only large taxonomic group restricted to a geographical area, roughly delimited by the N and S tropics. Overall, density increased regularly from cluster 1 (203 ± 16 individuals/m²) to clusters 4 (1,611 ± 150 individuals/m²) and 5 (1,662 ± 111 individuals/m²). Earthworms increased from cluster 1 to 3, then decreased in clusters 4 and 5, whereas Coleoptera, ants and Myriapoda culminated in cluster 4 and termites in cluster 5.

3.4 | Determinants of macroinvertebrate communities

Discriminant analyses showed highly significant ($p < .01$) effects of plant cover (11.8% of total variance explained), elevation (14.8%), latitude (12.8%), soil texture (3.9%) and temperature (8.7%) (Table S1). Highest abundance would be found at 1,000–2,000 m elevation, in equatorial (0–15°) latitudes, with rainfall ranging from 2,000 to 2,500 mm/year and average temperature in the 20–30°C range (Figure S5).

Plant cover had a significant effect on taxonomic richness (Figure S6), total density and community composition (Figure 4). The
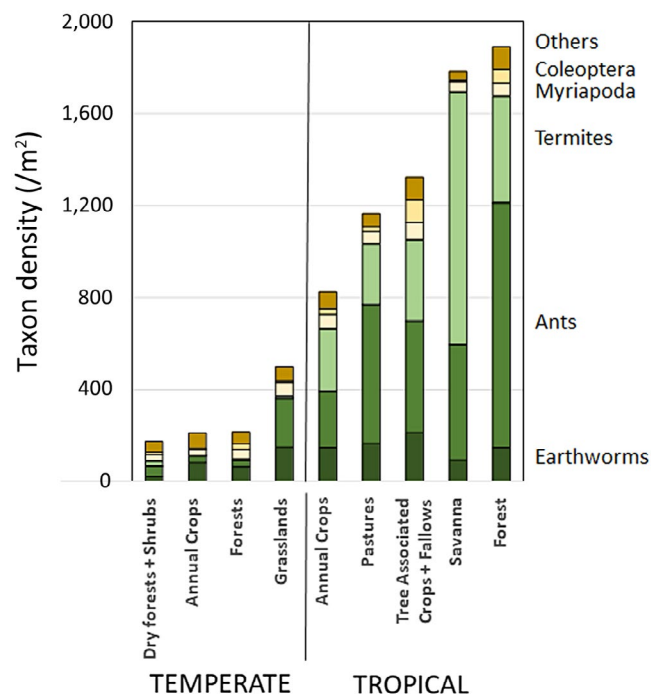
mean (±SE) number of orders found at a site varied from 3.1 ± 0.14 in dry forest and scrub vegetation sites to 9.3 ± 0.24 in tropical rain forests. (Figure S6). Plant covers in tropical areas exhibited taxonomic richness values higher than six, whereas those in temperate regions had values from four to five. The unexpectedly high number of outliers (shown in Figure S6) for temperate forests corresponds to subtropical deciduous forests located in Mexico at elevations of c. 1,500 m, and a few deciduous forests from France. A similar feature observed for the temperate dry forests and shrub vegetation corresponds to scrub vegetation at >3,000 m elevation, where a relatively low rainfall is compensated by low evaporation rate.

Overall community densities followed similar trends to taxonomic richness, with minimum densities of <200 individuals/m² in urban and peatbog soils, 200–250 individuals/m² in bare soils, annual crops from temperate areas and temperate forests, to maximum values of 1,895 ± 234 individuals/m² in tropical rain forests and 1,796 ± 72 individuals/m² in natural savannahs (Figure 4). Although tropical crops associated with tree components in agrosylvopastoral, agroforestry or tree plantations generally had rather high macroinvertebrate community densities, annual cropping systems with no trees had a highly detrimental influence with an overall decrease of c. 64% in temperate areas, if compared with temperate grasslands (but not different when compared with forests), and 54% when comparing tropical annual crops with the original rain forest.

3.5 | Covariation among macroinvertebrate communities and environmental parameters

Coinertia analysis between environmental and macroinvertebrate communities showed a significant covariation (coefficient of matrix correlation [RV] = 0.20; $p < .001$; Figure 5) and confirmed individual tests done on single categories of environment parameters. Tropical rain forests, tropical crops with tree components, temperate and tropical grasslands and high-rainfall conditions (>2,000 mm/year) were associated with high taxonomic richness and high densities of all macrofauna groups, with social insects (ants and termites) being particularly important, along with Coleoptera and earthworm densities and occurrence of the relatively rare "other" groups.

At the other end of the axis, sites with low densities and taxonomic richness were often those with bare soils or supporting annual crops and were often located in low-rainfall areas within subtropical areas in the Southern Hemisphere.

3.6 | Network analyses

We found generally higher numbers of significant associations when grouping communities according to plant cover than according to clusters (Figure 6). Dry and humid tropical forests exhibited maximum numbers of 73 and 79 interactions, respectively, from a maximum of 91 (Figure 6). Temperate and tropical grasslands and tropical tree-associated crops still showed rather high values, >60,
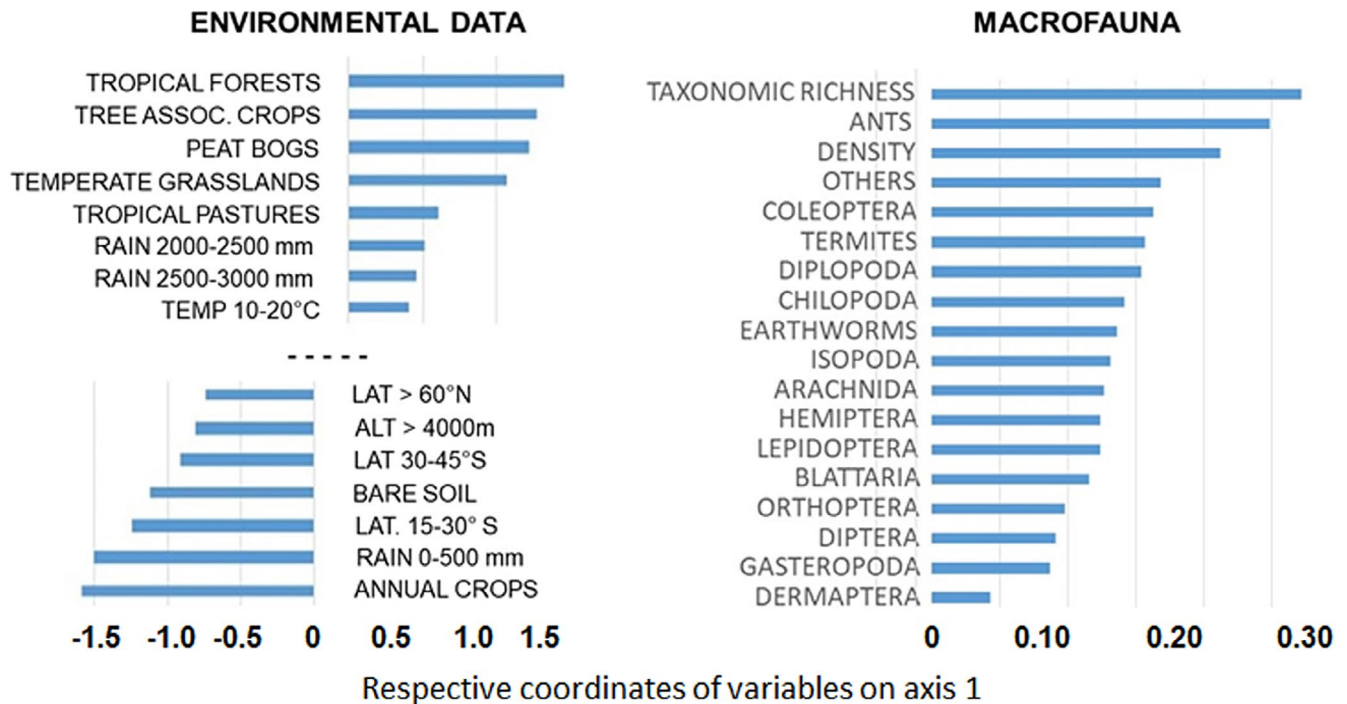

FIGURE 4 Compositions of macroinvertebrate communities in different types of plant cover considered in the 3,964 sites of the Macrofauna database

FIGURE 5 Respective coordinates of the 15 most influential environmental variables (of a total of five and their 41 modalities; left) and macroinvertebrate groupings (right) on axis 1 of a coinertia analysis among environmental and macroinvertebrate groupings of the 3,964 sites of the Macrofauna database. Groupings from both sides that have similar coordinates along factorial axes exhibit maximum covariation

whereas temperate forests had a rather low value of 37. Minimum values were observed in temperate and tropical annual crops, with 28 and 29 significant interactions, respectively.

Positive interactions prevailed except in temperate forests and annual crops, with maximum values in tropical forests (59), tropical pastures and temperate dry forests and shrubs (47) and temperate grasslands (41), whereas the largest occurrences of negative interactions (26) were observed in tropical forests and temperate grasslands and tropical tree-associated systems (24). Higher numbers of clusters and values of modularity in negative interactions show a more fragmented pattern in negative than positive interactions (Table 2).

Links among taxonomic groups measured by their co-occurrence also varied greatly among clusters (Figure S7). Cluster 1 is characterized by a high connectance: all groups are linked to at least another one, and 74.3% of the possible interactions are significant. Cluster 2 presents a similar pattern, with 36 positive and 19 negative connections. Clusters 4 and 5 exhibit intermediate patterns, whereas cluster 3 exhibits the lowest number of significant connections, with only 17 links, 21.8% of the possible total. Connectivity indices, which summarize this information, follow similar patterns, with the highest values for clusters 1 and 2 and lower values for clusters 3–5.

The rather reduced communities grouped in clusters 1 and 2 are highly structured, with high connectance and taxonomic groups related into only one cluster, whereas more populous communities from clusters 3–5 tend to consist of between 5 and 10 clusters of interconnected taxa instead of only one, as in the first two groups.

Modularity values vary among clusters, with the highest value found in the most developed communities of cluster 5 (Table S1).

Taxonomic groups show a large diversity of patterns of associations that might partly explain the differences observed among plant covers and clusters (Figure S8). The three main groups of ecosystem engineers exhibited mostly negative associations with other taxa. On the contrary, the litter transformers (Diplopoda, Isopoda and Chilopoda) had the highest number of positive associations with other taxa.

4 | DISCUSSION

Clear patterns in community composition and structure have been revealed through analysis of the present large data set and have confirmed the dominant role of plant cover and other environmental parameters in determining macroinvertebrate communities. We discuss here the general methodology used and the hypothesis testing proposed.

4.1 | Methods

Macroinvertebrate communities have been assessed across a wide range of sites using the simple low-cost ISO/TSBF method described here. Comparing data obtained from so many places must inevitably have been associated with differences in the quality of
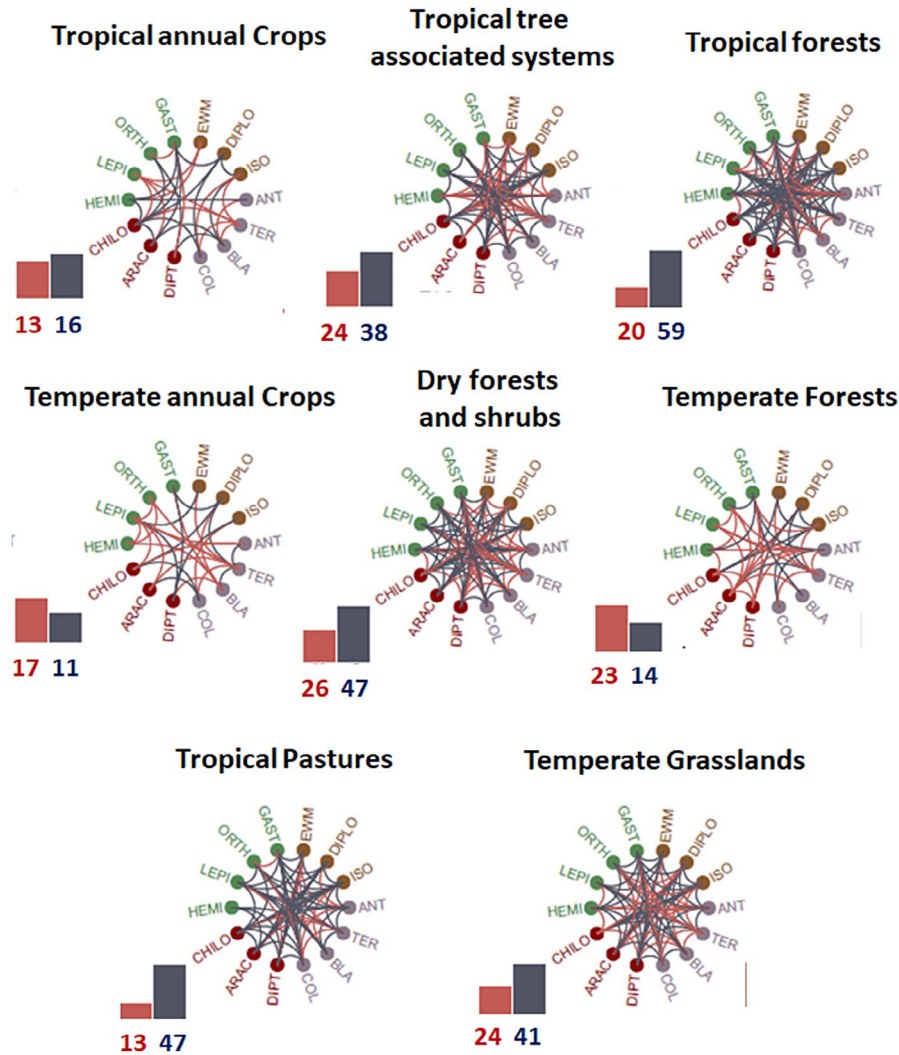
FIGURE 6 Network analysis of macroinvertebrate communities of the Macrofauna database in the eight main types of plant cover. Black bars indicate the number of significant positive relationships; red bars indicate negative relationships

TABLE 2 Network characteristics in groups of sites with similar plant cover

Cover	Positive connectivity	Positive clusters (n)	Positive modularity	Negative connectivity	Negative clusters (n)	Negative modularity
Tropical annual crops	0.55	3	0.22	0.21	4.00	0.40
Tree-associated tropical crops	0.46	2	0.23	0.10	6.00	0.47
Tropical forests	0.62	3	0.22	0.27	4.00	0.38
Temperate annual crops	0.50	3	0.24	0.32	3.00	0.29
Dry forests	0.47	3	0.22	0.13	6.00	0.54
Temperate forests	0.55	3	0.24	0.18	4.00	0.29
Temperate grasslands	0.59	3	0.21	0.27	3.00	0.29
Tropical pastures	0.65	3	0.13	0.23	4.00	0.43

the separation. However, the large number of different projects mostly led by people long trained in the method led us to expect a rather homogeneous level of accuracy. Invertebrates from so many different groups are not all captured with the same accuracy owing to differences in colour, mobility and size. As a result, the data obtained should be considered accurate indicators of the state of communities, not exact evaluations of the densities of all individual groups. In addition, high local variability in estimated densities was

to be expected from sampling such a large number of different taxonomic groups, with their often-aggregated distributions, and using relatively small sampling units (Ettema & Wardle, 2002; Jimenez et al., 2001). Patchy distributions frequently observed in soil macroinvertebrate populations probably explain the pattern observed in most taxonomic groups that have a large number of zero values and medians equal to zero (Table 1). Despite these potential limitations, analysis of the data set has revealed clear patterns. Both natural environmental determinants and human impacts on plant cover were assessed globally, and their significances and hierarchical arrangements were tested. Success in this area is probably attributable to the particular sensitivity of macroinvertebrate communities to environmental drivers and the robustness of this simple sampling method.

Other authors have analysed subsets of this and other databases, using levels of taxonomic resolution from orders to families and species. They concluded that orders and families have nearly the same capacity as species or genera to discriminate among the effects of the major environmental general drivers (Nahmani et al., 2006; Ruiz et al., 2011). This underlines the need to consider equivalent levels of scale resolution in comparing sites and taxonomic groups. Sites separated at the moderately broad spatial resolutions (a forest, a cropping plot) generally used in these studies are often better discriminated using coarse taxonomic units (orders and families) than by using lower levels, such as genera and species. In contrast, population dynamics studies or small-scale studies considering the effect of a single plant will always require methodologies specific to individual taxa (Moreira et al., 2008).

A possible drawback of the present method is that its efficiency has not yet been assessed clearly. Determining the extraction efficiency of macroinvertebrates from their litter and soil substrates depends on a number of factors: their size, colour, mobility and the competence of the operators. Based on the results of Lavelle et al. (1981), who checked the efficiency of a similar technique using much larger sampling units, we suspect that a general underestimation from 10 to 25% is likely to have occurred in many site-based studies. The procedure proposed by Velasquez (2020) offers a strict separation between groups and might allow recoveries close to 100% for most invertebrates.

4.2 | Diversity and distribution of communities

Our results confirm the constant presence, quantitative importance and diversity of macroinvertebrate communities world-wide. This finding supports our introductory statement that such an important biological component cannot be overlooked in conceptual and simulation models of soils. Cluster analysis has indicated that macroinvertebrate communities with comparable levels of diversity and abundance can occur across a broad range of geographical conditions, although this might not apply in extreme, cold or high-elevation environments. In addition, all groups, except termites,

could be found at all the sites, covering a wide range of soil and climate conditions. Although several significant environmental drivers have been identified, no taxa seem to be specific to a particular geographical area or plant cover type.

In general terms, systems that have a well-developed tree component and surface litter cover will host more diverse and abundant communities of litter transformers. They are Myriapoda, Isopoda, Arachnida and many other groups that can be excluded from pastures or cropped fields when the soil is kept bare for a large part of the year and/or when pesticides spread at the soil surface decimate their populations. In grasslands and herbaceous cropping systems, however, other groups, such as earthworms and Coleoptera, have the potential to become dominant. Although the conditions of these ecosystems, in particular specific water regimes and abundance of fine roots in soil, may favour them, the absence of trees and of the litter layer below them has detrimental effects on the whole litter-feeding community.

The difference in overall abundance between annual cropping systems deprived of trees and the largest community, −54% compared with forest in tropical areas and −64% compared with grasslands in temperate areas, sustains this statement. These global average values cover a wide range of variations that can be understood only by considering the detail of local and regional variations. In this analysis, for example, Diptera larvae appeared as a minor component of communities despite reported densities of several hundred individuals per square metre in a large number of European soils (Frouz, 1999). Although a great number of them are small and belong to the mesofauna that is not collected in the ISO/TSBF sampling, they also appear to be more abundant in specific biogeographical areas and ecosystems.

All the environmental factors tested (soil texture, temperature, rainfall, latitude and plant cover) impact soil communities and account for significant but variable percentages of overall variance, thus verifying our first hypothesis. The data set, however, did not allow us to test for their possible organization into the hierarchical suite of determinants, from climate to soil and organic matter quality, proposed by Lavelle et al., 1993).

An extended data set incorporating sites with larger ranges of climatic variables and more precisely defined definitions of soil conditions and litter quality (both surface and root litter) might result in a greater discrimination between macroinvertebrate communities. The present database lacks data from cold and hot, semi-desert and desert areas. We also need consistent definitions of soil taxonomic groupings to test the effect of soil age and a possible trade-off between young soils, rich in nutrients but shallow, and old soils that are deep and well structured, but are generally depleted of nutrients and organic matter. A preliminary analysis suggested that earthworms are more abundant in younger soils, whereas termites dominate in older soils, and ants are abundant in the youngest and oldest soils (Orgiazzi et al., 2016).

The amount of variance explained on each axis by individual environmental drivers is limited in all cases, although highly significant.

The highest value recorded for plant cover (12.3% of variance explained) probably expresses the strong overall effect of soil management, because >60% of the samples contained in the data base were derived from agricultural environments. Relationships with elevation (8.9%), latitude (5.3%), soil texture (3.9%) and temperature (2.3%) were also significant, although we do not yet know whether the differences found among these environmental factors might also be significant, allowing them to be classified by their importance.

4.3 | Vulnerability and reconstitution of communities

Many soil invertebrate species have become extinct, even before it was possible for them to be collected and described by taxonomists (Lavelle & Lapied, 2003). Earthworm communities lose species as a consequence of the conversion of pristine ecosystems to agricultural land. Peregrine species with world-wide distributions replace the many local species that are unable to compete successfully in such altered environments. Termite communities also suffer from simplification associated with deforestation, whereas ant communities seem better able to adapt to disturbances globally (Decaëns et al., 2018; Lawton et al., 1998).

A comparison of communities between the original tropical forest and derived cropping systems without trees shows an overall decrease of 54% in the abundance of macroinvertebrate communities and 25% in the average number of taxonomic units. This result allows us to verify our second hypothesis. Termites, ants and litter transformers seem to be the most affected. In contrast, earthworm populations might increase owing to colonization by invasive species, especially in tree-based cropping systems. In temperate regions, cropped systems have c. 64% lower densities than grasslands, but densities are almost equal to those in forests. The differences that occur between forests and grasslands might be a consequence of recent climatological changes. The last glaciation, some 25,000 years ago, left bare soils in the Northern Hemisphere that were initially covered with grasslands, which might have become the ecosystem where communities first recovered, rather than the forest that developed later. Mixing many allochthonous coniferous forests, which produce low-quality leaf litter, with natural deciduous forests might also explain part of this trend.

Beyond this general loss of biodiversity and abundance in agricultural systems, the high plasticity and reactivity of communities at local scales suggest that rapid recuperation is possible in ecosystem restoration operations or during shifts from conventional to organic agriculture. Examples have been described commonly from both temperate and tropical areas (Curry & Good, 1992; Mathieu et al., 2005; Velasquez et al., 2012). These restored communities might not have the same biodiversity and global functional profile as the original ecosystem, but their effects on such basic soil properties as macro-aggregation and soil-based ecosystem services are immediate (Velasquez & Lavelle, 2019).

4.4 | Connectedness

Analysis of networks in communities has shown that connectivity (the interdependence of populations) is maximal and modularity (the existence of independent groups of interacting populations) minimal in communities with reduced abundance and taxonomic richness. This contrasts with results of the study by Morriën et al. (2017), who showed that connectivity diminished with increasing soil disturbance. A possible explanation might be that highly developed communities (found where suitable conditions prevail) have high modularity. This means that they are divided into a number of subcommunities with closely associated species (sometimes rare species) or species combinations within specific small-scale conditions (Margerie et al., 2001). In disturbed systems at landscape to regional scales, communities are under stress, and homogenization has occurred. This leads to simpler communities, with large numbers of ubiquitous species and few of the most vulnerable specialist and rare species. The reasons why they become more connected has not yet been explored.

4.5 | Next step

The high sensitivity of macroinvertebrate communities to environmental conditions and the rapid changes that occur in field experiments (Velasquez et al., 2012) suggest that community-level analysis might provide particularly good indicators of soil quality. Significant associations between macroinvertebrate communities and soil (physical, chemical and morphological) data have been reported regularly (Lavelle et al., 2014; Rousseau et al., 2013; Velasquez & Lavelle, 2019). A further step would be to calibrate the general indicator extracted from macroinvertebrate communities against similar indicators of chemical, physical and morphological quality, looking for general patterns. Contrasting macroinvertebrate indicators derived from soils with sparse, low-diversity communities with those that are diverse and fully developed has shown that healthy soils always have higher macroinvertebrate indexes (Grimaldi et al., 2014; Lavelle et al., 2014).

An interesting further step will be to identify which level of taxonomic precision provides the best indicator values and to test the statement made by Nahmani et al. (2006) that large taxonomic units are better indicators than species.

The ultimate goal of this research should be to provide farmers and other soil managers with easy tools to assess soil quality and measure the impact of agricultural practices on the provision of soil-based ecosystem services. Field operators generally have a fairly good knowledge of the local macroinvertebrate communities and might readily be trained in sampling and identifying a small number of key indicator taxa.

Finally, plant community management is a key practice to develop in agroecological farming systems that would benefit from having active and diverse soil macroinvertebrate communities. The design of efficient associations of crops and companion plants requires a

clear understanding of the relationships between above- and below-ground communities, and faunal indicators could be used to orientate relevant methodological developments (Chave et al., 2014).

5 | CONCLUSIONS

As expected, macroinvertebrate communities were found to be very sensitive to climatic and soil conditions in their environment. However, surprisingly similar patterns of composition in terms of large taxonomic groups can occur in very different geographical areas. The five clusters identified all include sites from both temperate and tropical regions. This shows the importance of all environmental conditions in determining the abundance and diversity of macroinvertebrate communities and the consequences of agricultural practices. Although the present study emphasizes the general importance and plasticity of soil macroinvertebrate communities, their general absence in soil functioning models and agricultural practices is a serious obstacle to designing realistic modelling scenarios and forecasts (Bottinelli et al., 2015; Filser et al., 2016). Their fundamental roles as ecosystem engineers and key actors of self-organized soil processes need further clarification and recognition, well beyond the reductionist vision of their roles in energy and nutrient fluxes in food web processes promoted in a large number of publications (276 vs. 36 for ecosystem engineers) (Blouin et al., 2013; Folgarait, 1998; Jouquet et al., 2011). The overall functioning of the compartment of the litter-transforming populations and a more complete understanding of connectivity processes within the litter and soil systems need more research. Biochemical engineering that operates through the release of signal molecules by invertebrates and microorganisms within the soil increasingly appears to be another highly promising avenue of research (Chave et al., 2014). Modulation of plant gene expression via complex invertebrate–microorganism–root interactions is now widely recognized as a consequence of the self-organization mechanisms that occur in soils (Blouin et al., 2005; Jana et al., 2010).

A comprehensive understanding of underlying mechanisms will sustain the next generation of agricultural practices that will aim to optimize the management of soil biological potential (Blanchart & Trap, 2020). The development and use of general community profiles and indicator taxa by farmers is another priority goal for research. Such tools developed by scientists and other fieldworkers will allow measurement of the different ecosystem services provided by soils and the use of this knowledge to sustain novel public policies and thereby support the next generation of sustainable agricultural practices.

Beyond these immediate practical issues, the Macrofauna database will allow the inclusion of macroinvertebrate communities in global to local comparisons of distribution patterns of different groups of organisms, soil-dwelling microorganisms, micro- and mesofauna and plants. Hence, it is expected that different patterns will appear in the relative importances and functional profiles of the groups, determined by environmental conditions. Although general climatic conditions are always expected to predominate, general soil chemical and physical fertility and the quality of organic resources provided by plants should be important modifiers in final patterns.

ORCID

Patrick Lavelle https://orcid.org/0000-0002-2127-1067

REFERENCES

Atmar, W., & Patterson, B. D. (1993). The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, *96*(3), 373–382.

Barot, S., Blouin, M., Fontaine, S., Jouquet, P., Lata, J.-C., & Mathieu, J. (2007). A tale of four stories: Soil ecology, theory, evolution and the publication system. *PLoS One*, *2*(11), e1248. https://doi.org/10.1371/journal.pone.0001248

Blanchart, E., & Trap, J. (2020). Intensifier les fonctions écologiques du sol pour fournir durablement des services écosystémiques en agriculture 121. *Étude Et Gestion Des Sols*, *27*, 121–134.

Blouin, M., Zuily-Fodil, Y., Pham-Thi, A. T., Laffray, D., Reversat, G., Pando, A., Tondoh, J. E., & Lavelle, P. (2005). Belowground organism activities affect plant aboveground phenotype, inducing plant tolerance to parasites. *Ecology Letters*, *8*(2), 202–208.

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. *European Journal of Soil Science*, *64*(2), 161–182. https://doi.org/10.1111/ejss.12025

Bottinelli, N., Jouquet, P., Capowiez, Y., Podwojewski, P., Grimaldi, M., & Peng, X. (2015). Why is the influence of soil macrofauna on soil structure only considered by soil ecologists? *Soil & Tillage Research*, *146*, 118–124.

Brussaard, L., Pulleman, M. M., Ouedraogo, E., Mando, A., & Six, J. (2006). Soil fauna and soil function in the fabric of the food web. *Pedobiologia*, *50*(6), 447–462.

Cameron, E. K., Martins, I. S., Lavelle, P., Mathieu, J., Tedersoo, L., Bahram, M., Gottschall, F., Guerra, C. A., Hines, J., Patoine, G., Winter, M., Cesarz, S., Ferlian, O., Kreft, H., Lovejoy, T. E., Montanarella, L., Pereira, H. M., Phillips, H. R. P., Settele, J., … Eisenhauer, N. (2019). Global mismatches in aboveground and belowground biodiversity. *Conservation Biology*, *33*(5), 1187–1192.

Chave, M., Tchamitchian, M., & Ozier-Lafontaine, H. (2014). Agroecological engineering to biocontrol soil pests for crop health. In H. Ozier-Lafontaine, & M. Lesueur-Jannoyer (Eds.), *Sustainable agriculture reviews* (Vol. *14*, pp. 269–297). Springer International Publishing.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*. https://igraph.org

Curry, J. P., & Good, J. A. (1992). In Soil faunal degradation and restoration. In R. Lal, & B. A. Stewart (Eds.), *Soil restoration* (pp. 171–215). Springer.

Decaëns, T. (2010). Macroecological patterns in soil communities. *Global Ecology & Biogeography*, *19*, 287–302.

Decaëns, T., Martins, M. B., Feijoo, A., Oszwald, J., Doledec, S., Mathieu, J., de Sartre, X. A., Bonilla, D., Brown, G. G., Criollo, Y. A. C., Dubs, F., Furtado, I. S., Gond, V., Gordillo, E., Le Clec'h, S., Marichal, R., Mitja, D., de Souza, I. M., Praxedes, C., … Lavelle, P. (2018). Biodiversity loss along a gradient of deforestation in Amazonian agricultural landscapes. *Conservation Biology*, 32(6), 1380–1391.

Doledec, S., & Chessel, D. (1994). Co-inertia analysis–An alternative method for studying species environment relationships. *Freshwater Biology*, 31(3), 277–294.

Dray, S., & Dufour, A. B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22(4), 1–20. https://doi.org/10.18637/jss.v022.i04

Ettema, C. H., & Wardle, D. A. (2002). Spatial soil ecology. *Trends in Ecology & Evolution*, 17(4), 177–183. https://doi.org/10.1016/S0169-5347(02)02496-5

Filser, J., Faber, J. H., Tiunov, A. V., Brussaard, L., Frouz, J., De Deyn, G., Uvarov, A. V., Berg, M., Berge, P., Lavelle, P., Loreau, M., Wall, D. H., Quernerio, P., Eijsackers, H., & Jimenez, J. J. (2016). Soil fauna: key to new carbon models. *Soil*, 2(4), 565–582. https://doi.org/10.5194/soil-2-565-2016

Folgarait, P. J. (1998). Ant biodiversity and its relationship to ecosystem functioning: A review. *Biodiversity & Conservation*, 7(9), 1221–1244.

Frouz, J. (1999). Use of soil dwelling Diptera (Insecta, Diptera) as bioindicators: A review of ecological requirements and response to disturbance. *Agriculture Ecosystems & Environment*, 74(1–3), 167–186.

Grimaldi, M., Oszwald, J., Dolédec, S., Hurtado, M. D. P., de Souza Miranda, I., Arnauld de Sartre, X., Assis, W. S. D., Castañeda, E., Desjardins, T., Dubs, F., Guevara, E., Gond, V., Lima, T. T. S., Marichal, R., Michelotti, F., Mitja, D., Noronha, N. C., Delgado Oliveira, M. N., Ramirez, B., … Lavelle, P. (2014). Ecosystem services of regulation and support in Amazonian pioneer fronts: Searching for landscape drivers. *Landscape Ecology*, 29(2), 311–328. https://doi.org/10.1007/s10980-013-9981-y

Guerra, C. A., Bardgett, R., Caon, L., Crowther, T. W., Delgado-Baquerizo, M., Montanarella, L., Navarro, L. M., Orgiazzi, A., Singh, B. K., Tedersoo, L., Vargas-Rojas, R., Briones, M. J., Buscot, F., Cameron, E. K., Cesarz, S., Chatzinotas, A., Cowan, D. A., Djukic, I., van den Hoogen, J., … Eisenhauer, N. (2021). Tracking, targeting, and conserving soil biodiversity. *Science* 1, 371, 6526, 239–241. https://doi.org/10.1126/science.abd7926

ISO. (2011). *ISO 23611–5. Soil quality-sampling of soil invertebrates-Part 5. Sampling and extraction of soil macro-invertebrates*. International Organization for Standardization.

Jana, U., Barot, S., Blouin, M., Lavelle, P., Laffray, D., & Repellin, A. (2010). Earthworms influence the production of above- and belowground biomass and the expression of genes involved in cell proliferation and stress responses in Arabidopsis thaliana. *Soil Biology & Biochemistry*, 42(2), 244–252.

Jiménez, J. J., Rossi, J.-P., & Lavelle, P. (2001). Spatial distribution of earthworms in acid-soil savannas of the eastern plains of Colombia. *Applied Soil Ecology*, 17(3), 267–278.

Jiménez, J. J., Lavelle, P., & Decaëns, T. (2006). The efficiency of soil hand-sorting in assessing the abundance and biomass of earthworm communities. Its usefulness in population dynamics and cohort analysis studies. Special issue XIVth International Colloquium on Soil Zoology (Rouen, France). *European Journal of Soil Biology*, 42(Suppl 1), S225–S230.

Jouquet, P., Traore, S., Choosai, C., Hartmann, C., & Bignell, D. (2011). Influence of termites on ecosystem functioning. Ecosystem services provided by termites. *European Journal of Soil Biology*, 47(4), 215–222. https://doi.org/10.1016/j.ejsobi.2011

Lavelle, P., Maury, M. E., & Serrano, V. (1981). Estudio cuantitativo de la fauna del suelo en la region de Laguna Verde, Vera-Cruz. Epoca de lluvias. *Instituto De Ecologia Publications. (Mexico)*, 6, 75–105.

Lavelle, P., Blanchart, E., Martin, A., Martin, S., Spain, A., Toutain, F., Barois, I., & Schaefer, R. (1993). A hierarchical model for decomposition in terrestrial ecosystems—Application to soils of the humid tropics. *Biotropica*, 25(2), 130–150.

Lavelle, P., & Lapied, E. (2003). Endangered earthworms of Amazonia: An homage to Gilberto Righi. *Pedobiologia*, 47(5–6), 419–427.

Lavelle, P., Rodriguez, N., Arguello, O., Bernal, J., Botero, C., Chaparro, P., Gomez, Y., Gutierrez, A., Hurtado, M. D., Loaiza, S., Pullido, S. X., Rodriguez, E., Sanabria, C., Velasquez, E., & Fonte, S. J. (2014). Soil ecosystem services and land use in the rapidly changing Orinoco River Basin of Colombia. *Agriculture Ecosystems & Environment*, 185, 106–117. https://doi.org/10.1016/j.agee.2013.12.020

Lavelle, P., Spain, A., Blouin, M., Brown, G., Decaëns, T., Grimaldi, M., Jiménez, J. J., McKey, D., Mathieu, J., Velasquez, E., & Zangerlé, A. (2016). Ecosystem engineers in a self-organized soil: A review of concepts and future research questions. *Soil Science*, 181(3–4), 91–109. https://doi.org/10.1097/ss.0000000000000155

Lavelle, P., Mathieu, J., Spain, A., Brown, G., Fragoso, C., Lapied, E., De Aquino, A., Barois, I., Barrios, E., Barros, M. E., Bedano, J. C., Blanchart, E., Caulfield, M., Chagueza, Y., Dai, J., Decaëns, T., Dominguez, A., Dominguez, Y., Feijoo, A., … Zhang, C. (2022). A worldwide assessment of soil macroinvertebrate communities. *Dryad*. https://doi.org/10.5061/dryad.r2280gbc2

Lawton, J. H., Bignell, D. E., Bolton, B., Bloemers, G. F., Eggleton, P., Hammond, P. M., Hodda, M., Holt, R. D., Larsen, T. B., Mawdsley, N. A., & Stork, N. E. (1998). Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature*, 391(6662), 72–76.

Margerie, P., Decaëns, T., Bureau, F., & Alard, D. (2001). Spatial distribution of earthworm species assemblages in a chalky slope of the Seine Valley (Normandy, France). *European Journal of Soil Biology*, 37, 291–296.

Mathieu, J., Rossi, J. P., Mora, P., Lavelle, P., Martins, P. F. D., Rouland, C., & Grimaldi, M. (2005). Recovery of soil macrofauna communities after forest clearance in Eastern Amazonia, Brazil. *Conservation Biology*, 19(5), 1598–1605.

Medina-Sauza, R. M., Alvarez-Jimenez, M., Delhal, A., Reverchon, F., Blouin, M., Guerrero-Analco, J. A., Cerdan, C. R., Guevara, R., Villain, L., & Barois, I. (2019). Earthworms building up soil microbiota, a review. *Frontiers in Environmental Science*, 7, 1–20. https://doi.org/10.3389/fenvs.2019.00081

Moreira, F. M. S., Bignell, D. E., & Huising, E. J. (2008). *A handbook of tropical soil biology: Sampling and characterization of below-ground biodiversity*. Earthscan, 205 p.

Morriën, E., Hannula, S. E., Snoek, L. B., Helmsing, N. R., Zweers, H., de Hollander, M., Soto, R. L., Bouffaud, M.-L., Buée, M., Dimmers, W., Duyts, H., Geisen, S., Girlanda, M., Griffiths, R. I., Jørgensen, H.-B., Jensen, J., Plassart, P., Redecker, D., Schmelz, R. M., … van der Putten, W. H. (2017). Soil networks become more connected and take up more carbon as nature restoration progresses. *Nature Communications*, 8, 14349.

Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B. J., Peet, R. K., Ott, J. E., Violle, C., Enquist, B. J., Jorgensen, P. M., & Svenning, J. C. (2016). A network approach for inferring species associations from co-occurrence data. *Ecography*, 39(12), 1139–1150.

Nahmani, J., Lavelle, P., & Rossi, J. P. (2006). Does changing the taxonomical resolution alter the value of soil macroinvertebrates as bioindicators of metal pollution? *Soil Biology and Biochemistry*, 38(2), 385–389.

Nielsen, U. N., Wall, D. H., & Six, J. (2015). Soil biodiversity and the environment. *Annual Review of Environment and Resources*, 40, 63–90.

Orgiazzi, A., Bardgett, R. D., Barrios, E., Behan-Pelletier, V., Briones, M. J. I., Chotte, J. L., De Deyn, G. B., Eggleton, P., Fierer, N., Fraser, T., Hedlund, K., Jeffrey, S., Johnson, N. C., Jones, A., Kandeler, E., Kaneko, N., Lavelle, P., Lemanceau, P., Miko, L., … Wall, D. H. (2016). *Global soil biodiversity atlas*. 176 p. European Commission, Publications Office of the European Union.

Perry, D. A. (1995). *Self-organizing systems across scales. Tree, 10,* 241–245.

Rousseau, L., Fonte, S. J., Tellez, O., van der Hoek, R., & Lavelle, P. (2013). Soil macrofauna as indicators of soil quality and land use impacts in smallholder agroecosystems of western Nicaragua. *Ecological Indicators, 27,* 71–82. https://doi.org/10.1016/j.ecolind.2012.11.020

Ruiz, N., Lavelle, P., & Jimenez, J. J. (2008). *Soil macrofauna field manual.* FAO (101 pp.).

Ruiz, N., Mathieu, J., Célini, L., Rollard, C., Hommay, G., Iorio, E., & Lavelle, P. (2011). IBQS: A synthetic index of soil quality based on soil macro-invertebrate communities. *Soil Biology & Biochemistry, 43*(10), 2032–2045.

Thebault, E., & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science, 329*(5993), 853–856.

Tsiafouli, M. A., Thébault, E., Sgardelis, S. P., de Ruiter, P. C., van der Putten, W. H., Birkhofer, K., Hemerik, L., de Vries, F. T., Bardgett, R. D., Brady, M. V., Bjornlund, L., Jørgensen, H. B., Christensen, S., Hertefeldt, T. D'., Hotes, S., Gera Hol, W. H., Frouz, J., Liiri, M., Mortimer, S. R., … Hedlund, K. (2015). Intensive agriculture reduces soil biodiversity across Europe. *Global Change Biology, 21*(2), 973–985.

Velasquez, E., Lavelle, P., & Andrade, M. (2007). GISQ, a multifunctional indicator of soil quality. *Soil Biology & Biochemistry, 39*(12), 3066–3080. https://doi.org/10.1016/j.soilbio.2007.06.013

Velasquez, E., Fonte, S., Barot, S., Grimaldi, M., Desjardins, T., & Lavelle, P. (2012). Soil macrofauna-mediated impacts of plant species composition on soil functioning in Amazonian pastures. *Applied Soil Ecology, 56,* 43–50.

Velasquez, E., & Lavelle, P. (2019). Soil macrofauna as an indicator for evaluating soil based ecosystem services in agricultural landscapes. *Acta Oecologica, 100,* 103446. https://doi.org/10.1016/j.actao.2019.103446

Velasquez, E. (2020). *Manejo integrado de plagas y enfermedades en platano y banano.* https://www.youtube.com/watch?v=BZHbNLMpLRs&t=476s

BIOSKETCH

Patrick Lavelle is an emeritus professor of the Paris Sorbonne University with a PhD in Ecology. He is a corresponding member of the French Academy of Science and was awarded the Prize of the French Ecological Society in 2018. He created and for 20 years directed the Tropical Soils Ecology Laboratory, a mixed unit between Paris University and Institut de Recherche pour le Développement. A specialist in the ecology of earthworms and other soil macroinvertebrates, he has developed methodologies to assess their abundances and roles in tropical soils and supervised >30 PhD theses in 15 different tropical countries from America, Africa and Asia. Author of 218 papers published in peer-reviewed journals, he is a co-author with Alister Spain of a Soil Ecology textbook published in 2001 by Kluwer editions.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Lavelle, P., Mathieu J., Spain A., Brown G., Fragoso C., Lapied E., De Aquino A., Barois I., Barrios E., Barros M. E., Bedano J. C., Blanchart E., Caulfield M., Chagueza Y., Dai J., Decaëns T., Dominguez A., Dominguez Y., Feijoo A., … Supp S. (2022). Soil macroinvertebrate communities: A world-wide assessment. *Global Ecology and Biogeography, 00,* 1–16. https://doi.org/10.1111/geb.13492