SOILGUARD

Sustainable soil management to unleash soil biodiversity potential and increase environmental, economic and social well-being

Grant Agreement no. 101000371

2.3 Impact of soil degradation and management on soil biodiversity

Work package	WP2 Soil Biodiversity status and soil multifunctionality			
Author(s)	Data curation and analysis, synthesis: Laura García-Velázquez (UA), Santiago Soliveres Codina (UA), Anna Clocchiatti (UvA), Pablo Sánchez Cueto (LEITAT), Giulia Bongiorno (WUR), Ron de Goede (WUR), Martin Hartmann (ETH), Elly Morriën (UvA), Franciska de Vries (UvA) In situ information, local sampling, specific information used in the report: Ina Alsina (LLU), Paula Barral (INTA), Joseph Blaise (FSUY), Helle Hestbjerg (DTI), Mathilde Jeanbille (INRAE), Zoka Melpomeni (NOA), Luis Daniel Olivares (UMH), Taina Pennanen (LUKE), Kerry Ryan (TEAGASC), Rochana Tangkoonboribun (TISTR), Tomas van de Sande (INAGRO), Toth Zoltan (MATE) Data repository: Panagiotis Vlacheas (WINGS)			
Reviewer(s)	Salvador Lladó Fernández (University of Barcelona) Briony Jones (UKCEH), Laurence Jones (UKCEH) Kostadin Atanasov (LEITAT), Cristina Yacoub López (LEITAT)			
Deliverable nature	Technical Report			
Dissemination level (Confidentiality)	Public			
Date of delivery	Contractual	31/05/2024	Actual	30/08/2024
Version	V 4.0			
Total number of pages	112			
Keywords	soil biodiversity, soil degradation, agriculture management, landscapes, relative importance index			

Document History

Issue Date	Version	Action
22/04/2024	V.1	First draft for internal review
08/05/2024	V.2	Revised draft for internal review
29/05/2024	V.3	Final version for submission
30/08/2024	V.4	Second revised version for submission

Contents

List of Figures

LDA scores were represented by arrows. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.

Figure 9. *Beta-diversity Cameroon soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been colored and shaped by the combination of Management and Degradation. Samples have been colored and shaped by the combination of Management and Degradation. Different sample size was obtained depend on the target taxa → Prokaryotes: CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5; Fungi: CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5; Protists: CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5; Nematodes: CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5; Annelida: these were discarded due to number of samples < 50% of the total (n=4). Micro-arthropods: these were discarded due to number of samples < 50% of the total (n=9). Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

Figure 10. *Beta-diversity Southern Ireland grassland soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been coloured and shaped by the combination of Management and Degradation. Different sample size was obtained depend on the target taxa → Prokaryotes: MONO.Low=5, MONO.Medium=5, MIXED.low=5, MIXED.Medium=5; Fungi: MONO.Low=5, MONO.Medium=5, MIXED.low=4, MIXED.Medium=5; Protists: MONO.Low=5, MONO.Medium=5, MIXED.Low=5, MIXED.Medium=5; Nematodes: MONO.Low=5, MONO.Medium=5, MIXED.low=5, MIXED.Medium =5; Annelida: MONO.Low=5, MONO.Medium=4, MIXED.low=4, MIXED.Medium=5. Micro-arthropods: MONO.Low=4, MONO.Medium=5, MIXED.low=4, MIXED.Medium=5.Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

Figure 11. *Beta-diversity Southern Ireland agroforestry soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been coloured and shaped by the combination of Management and Degradation. Different sample size was obtained depend on the target taxa → Prokaryotes: CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium=3; Fungi: CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium=3; Protists: CTRL.Low=2, CTRL.Medium=3, AGROF.Low=2, AGROF.Medium=3; Nematodes: CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium =3; Annelida: CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium=1. Micro-arthropods: CTRL.Low=2, CTRL.Medium=2, AGROF.low=1, AGROF.Medium=2.Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

Figure 12*. Beta-diversity West Finland forest soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been coloured and shaped by Management. Different sample size was obtained depend on the target taxa → Prokaryotes: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4; Fungi: FOLD=4, FMID=4, FYM=3, CCR=4, CCGA=3, CCS=4 Protists: FOLD=4, FMID=4, FYM=3, CCR=4, CCGA=1, CCS=3; Nematodes: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4; Annelida: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4; Micro-arthropods: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

Figure 13. *Co-occurrence networks in alternatively and conventionally managed soils of West Flanders (A, n=10 and B, n=10), Murcia (C, n=10 and D, n=10) and Latvia (E, n=10 and F, n=10). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists, Nematodes, Microarthropods and Annelids) and by morphological characterization (Nematodes, Microarthropods) from each site.*

Figure 14. *Co-occurrence networks in alternatively and conventionally managed soils of South Transdanubia (A,B, n=10), and fields of Middle Jutland/South Denmark managed alternatively (C), recently converted to alternative management (D, n=10) and conventionally managed (E, n=10). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists, Nematodes, Microarthropods and Annelids) and by morphological characterization (Nematodes, Microarthropods) for each site.*

Figure 15. *Co-occurrence networks in alternatively and conventionally managed soils of Buenos Aires (A, n=10 and B, n=10), West Cameroon (C, n=10 and D, n=10) and Chiangrai (E, n=14 and F, n=14). Strong positive correlations (rho >*

0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists and Annelids) for each site.

Figure 16. *Co-occurrence networks in alternatively and conventionally managed soils of Ireland pastures with mixture and monoculture vegetation (A, n=10 and B, n=10 and), Ireland grasslands with agroforestry and the relative controls (C, n=10 and D, n=10), as well as forests managed as continuous cover vegetation and with clear-cutting (E, n=12 and F, n=12). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists, Nematodes, Microarthropods and Annelids) and by morphological characterization (Nematodes, Microarthropods) from for each site.*

Figure 17. *Network metrics values and standard error represented for each NUTS-2 region - West Flanders (BE), Murcia (SP), Latvia (La), South Transdanubia (HU), Middle Jutland/Siddenmark (South Denmark)(DE), Buenos Aires (ARG), West Cameroon (CA), Chiangrai (TH), Southern Ireland pastures (IR PAS) and agroforestry sites (IR AG) and West Finland (F) - and management type. The metrics represented are: the number of edges of a network, connectance, modularity, transitivity and assortativity. Standard errors were calculated with 95% confidence using a null model distribution for that metric, based on 1000 permutations of the observed network. Management is indicated on the x-axis as ALT (alternative management), CON (conventional management), ALT OLD and ALT YOUNG (old and recent alternatively managed soils), MIX (mixture grassland vegetation), MONO (grassland monoculture), AG and AG CONTROL (agroforestry grassland and its control).*

Figure 18. *The frequency distribution of the Sorensen similarity index of nematode family taxonomic composition for the European SOILGUARD samples analysed microscopically and molecularly (eDNA) using two different primers (18SV4, 18SV6V8). The European SoilGuard samples included Spain, Hungary, Denmark, Hungary, Ireland, Latvia, Belgium and Finland, and 151 samples were available after rarefaction of the molecular data obtained with the 18SV4 primer, while 163 samples were available after rarefaction of the molecular data obtained with the 18SV6V8 primer. Left above panel: Similarity index for the presence-absence data for 18SV4 versus 18SV6V8 primers data (n=151), right above panel: Similarity index for the 18SV4 primer data versus microscope data (n=151), and left below: Similarity index for the 18SV6V8 primer data and microscope data (n=163).*

Figure 19. *The frequency distribution of the Sorensen similarity index of micro-arthropod family taxonomic composition for the European SOILGUARD samples analysed microscopically and molecularly (eDNA) using two different primers (18SV4, 18SV6V8). The European SOILGUARD samples included Spain, Hungary, Denmark, Hungary, Ireland, Latvia, Belgium and Finland, and 75 samples were available after rarefaction of the molecular data obtained with the 18SV4 primer, while 131 samples were available after rarefaction of the molecular data obtained with the 18SV6V8 primer. Left above panel: Similarity index for the presence-absence data for 18SV4 versus 18SV6V8 primers data (n=75), right above panel: Similarity index for the 18SV4 primer data versus microscope data (n=75), and left below panel: Similarity index for the 18SV6V8 primer data and microscope data (n=131).*

Figure 20. *Correlation between the Acari and Collembola family richness both calculated based on microscopic observations in the 164 SOILGUARD European sites (Spain, Hungary, Denmark, Hungary, Ireland, Latvia, Belgium and Finland). The sites with >15 Acari families were all sampled in Finnish forest soils.*

Figure 21. *Relationship between the nematode-based indicator Maturity Index obtained by the NINJA program by using the count table with taxonomic information based on the morphological analysis, and the count table table with taxonomic information based on the data obtained with the primer 18SV6V8. The r squared indicates the coefficient of determination, determining the proportion of variariance that is explained by the explanatory variable.*

Figure 22. *Biodiversity indicators correlation heatmap. Cells were coloured based on the correlation matrix, with blue gradient when closer to +1 and red when closer to -1, with correlations higher than 0.5, are coloured in blue when closer to +1 and red when closer to -1. Low correlations (r> |0.5|) were removed and represented as blank cells.*

Figure 23*. Beta-diversity Principal Coordinate Analysis (PCoA) calculated from most prevalent species (common species) of the sequencing datasets, across the cross-biome network sites belonged to croplands from EU regions. Significant correlated ASV and biodiversity or environmental variable (r > |0.5|; p.adjusted fdr < 0.1) with PCO scores were represented by black and red arrows, respectively. Abbreviations Supplementary Table S4 in Annex III.*

List of Tables

of acronyms are the institutions abbreviations, as commonly used in Soilguard. Different shading is added for visualisation purposes (to differentiate between different NCPs).

Abbreviations

Abbreviation	Meaning
WP	Work package
EU	European Union
NCP	Nature Contribution to People
SOC	Soil organic carbon
PLFA/NLFA	Phospholipid-derived/neutral derived fatty acids
N	Nitrogen
P	Phosphorus
C	Carbon
q-PCR	Quantitative PCR
COI	Cytochrome Oxidase I (International Barcode of Life Consortium)
AP	Annual precipitation
MAT	Mean annual temperature
PET	Potential annual evapotranspiration
AI	Aridity index
DE	Denmark
BE	Belgium
IE	Ireland
SP	Spain
CA	Cameroon
ARG	Argentina
TH	Thailand
HU	Hungary
LV	Latvia
FI	Finland
RII	Resistance Interaction Index
AIC	Akaike Information Criterion
NEI	Naturalness evaluation index
SWF	Small wood features
CAP	Canonical analysis of principal coordinates
ASV	Amplicon Sequence Variant
LDA	Linear Discriminant Analysis
FOLD	Forest clearcut old
FMID	Forest clearcut middle
FYM	Forest clearcut mounding
CCR	Forest continuous retained
CCGA	Forest continuus gap
CCS	Forest continuous single
ALT	Alternative management
CONV	Conventional management
IR PAS	Ireland pastures
IR AG	Ireland agroforestry
AMF	Arbuscular mycorrhizal fungi
TMSB	Total microbial biomass
AOB	Ammonia oxidizing bacteria
AOA	Ammonia oxidizing archaea
COMA_A	Complete ammonia oxydizing Nitrospira clade A quantification
COMA_B	Complete ammonia oxydizing Nitrospira clade B quantification
PDM	Precipitation of the driest month
PDQ	Precipitation of the driest quarter
PCQ	Precipitation of the coldest quarter
TSE	Temperature seasonality

MTDQ	Mean temperature of the driest quarter
CV	Coeficient of variation
NDVI	Normalized Difference Vegetation Index
MUF	4-Methylumbelliferyl
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
WHC	Water holding capacity
TAN	Total available nitrogen
PNF	p-nitrophenol
TBI	Tea bag index
AVAP	Available phosphorus
AMO	Potential ammonification rate
DEP	Potential depolymerization rate
NTR	Potential nitrogen transformation rate
BG	Activity of b-glucosidase
XYL	Activity of Xylanase
PHOS	Activity of phosphatase
NAG	Activity of b-N-acetylglucosaminidase
CONDUC	Conductivity
INFILT	Infiltration
AGGR	Aggregates
LD	Litter decomposition
LDF	Leaf damage fungi
LDH	Leaf damage herbivores
N RET	Nitrogen retained by soil (lixiviates)
P RET	Phosphorus retained by soil (lixiviates)
SD	Soil degradation
MAN	Soil management
BIO	Biodiversity (index)

Summary

Soils are one of the largest reservoirs of biodiversity, which is increasingly threatened by soil degradation, unsustainable soil management and climate change. Despite the importance of monitoring soil, we have limited knowledge about the impact of these stressors on the status of soil biodiversity. Additionally, we are unaware of their cascading effects on soil multifunctionality and therefore, its capacity to maintain the delivery of nature's contributions to people.

To evaluate the impact of soil degradation and soil management on soil biodiversity, and also their cascading effects on soil multifunctionality, SOILGUARD conducted a comprehensive soil sampling across the cross-biome network of sites created in D2.1. The cross-biome network of sites includes ten EU NUTS-2 and international regions: Buenos Aires (Argentina), West Flanders (Belgium), West Cameroon, Middle Jutland/South Denmark (Denmark), South Transdanubia (Hungary), Latvia, Murcia (Spain), Chiangrai (Thailand), Southern Ireland, and West Finland. It covers eight distinct regions, including Atlantic, continental, Pannonian, Mediterranean, boreal, tropical humid, tropical savannah, and temperate oceanic, as well as three different land use categories: cropland, grassland, and forest, including different levels of soil degradation.

In this deliverable we present the advances of our project in terms of evaluating the main impacts of soil management and degradation on soil biodiversity and how this, in turn, can determine soil multifunctionality. In particular, we analysed the relative impact of agricultural management strategies (conventional versus alternative) on soil biodiversity and multifunctionality. We also analysed the influence of soil biodiversity on the delivery of multiple soil functions. Additionally, we designed artificial landscapes to determine the percentage of alternative/conventional agriculture necessary to maximize soil biodiversity and soil multifunctionality. The results presented here lay the groundwork for further analyses -within and beyond SOILGUARD- that will focus on determining the key management strategies for soil sustainability and will enable informed decision-making on the implementation of the EU directives related to the EU Green Deal. Lastly, as project results were obtained, we proposed as an additional objective the search for monetary cost-effective indicators of soil biodiversity.

Although still preliminary, the monitoring of three indicators (i.e., total microbial biomass, target genes related with the nitrogen cycle, and microarthropod abundance) could represent a good approach for the assessment of soil biodiversity status, as these encompass different "collections" of indicators showing contrasting environmental responses. Also, the results from this report underscore the importance of transitioning to organic farming practices for optimal sustainability outcomes across agrosystems' key dimensions: biodiversity, multifunctionality, and crop yield with a recommended increase of up to 50% on the total European agricultural land under organic management. The implementation of organic agriculture would be more beneficial if implemented at moderately to highly degraded soils, as it is under those conditions were beneficial effects are most evident for soil biodiversity and soil multifunctionality.

1. Introduction

Using soil degradation maps as a foundation, the SOILGUARD cross-biome network of sites was established, identifying sampling sites with both conventional and alternative soil management across at least two tiers of soil degradation in each region (as outlined in Deliverable 2.1, "Cross-biome network of sites set up"). Subsequently, in deliverable 2.2, an evaluation of soil biodiversity status was conducted across European NUTS-2 and global regions. Thus, the main objective of this deliverable (2.3) is to determine the impact of soil degradation and soil management on soil biodiversity and its potential to deliver soil multifunctionality. This will enable SOILGUARD to generate recommendations that can guide monitoring, evaluation, and the future development of sustainable soil management practices. To achieve this, we conducted a series of analyses that fulfilled the tasks of quantifying the effects of soil degradation and soil management strategies on soil biodiversity (Task 2.3). Additionally, we defined the role of soil degradation and soil management strategies in the conservation or loss of multifunctionality and quantified the weight of soil biodiversity in the delivery of each of the NCP assessed (Task 2.4). To achieve the objectives of Task 2.4, we addressed different subtasks based on the determination of soil multifunctionality, the relationships between soil biodiversity and multifunctionality, and the region-specific effects of soil management and soil degradation on soil biodiversity and soil multifunctionality.

The SOILGUARD project included organically managed farmland from different countries, both within the EU and from other continents (described in detail in D2.2 and 3.3). Depending on the region, the regulations for establishing a field as organic vary in stringency, with specific requirements and timelines for obtaining organic certification differing even within the EU. It is important to note that the definition of organic farming can sometimes be ambiguous. Therefore, the lowest common denominator, and what we classified as organic farming in SOILGUARD, was soil management that is free from chemical fertilisers and prohibited pesticides according to organic regulations.Other types of sustainable soil management practices such as crop rotation, cover cropping or conservational tilling, are far beyond organic agriculture and not included in the Farm to Fork Strategy, which only states as objective at least 25% of organic agriculture for 2030. For grasslands and forests, these included mixed vegetation (polycultures, vs monocultures) and continuous-cover (vs clear-cut) forests, commonly considered as more sustainable practices in the literature.

2. Summary of previous steps

Briefly, the cross-biome network of sites includes 233 locations, distributed in ten regions, seven of which are located in Europe NUTS-2 and three internationally (Figure 1). Buenos Aires (Temperate oceanic region), West Flanders (Atlantic region), West Cameroon (Tropical humid region), South Transdanubia (Pannonian region), Latvia (Boreal region) and Murcia (Mediterranean region) featured 20 sampling sites in croplands, while Chiangrai (tropical savannah region) and Middle Jutland/South Denmark (continental region) featured 30 and 29 cropland sites, respectively, as they included two periods since shifting to alternative management (Denmark) or three levels of soil degradation instead of two (Thailand). In total, croplands counted 179 sites. West Finland (Boreal region) had 24 forest sites and Southern Ireland (Atlantic region) had 20 grassland sites and 10 agroforestry sites (Deliverable 2.1, Map of the cross-biome network of sites and soil degradation gradients).

The study sites covered a wide variety abiotic factors (e.g., soil type, climate) and crop types including vegetables (potatoes) and cereals (maize, wheat, oatmeal, barley, wheat, etc). Annual precipitation (AP), annual mean temperature (MAT), annual potential evapotranspiration (PET) and aridity index (AI, United Nations Environment Programme, 1992) for each sampling site were extracted from WorldCLIM 2.0 (Fick and Hijmans, 2017) and Global Aridity and PET (CGIAR-CSI) datasets (Trabucco and Zomer, 2018). AP and MAT ranged from 295 mm/year to 1995 mm/year and 5.2 ^{o}C to 27.9 ^{o}C, respectively.

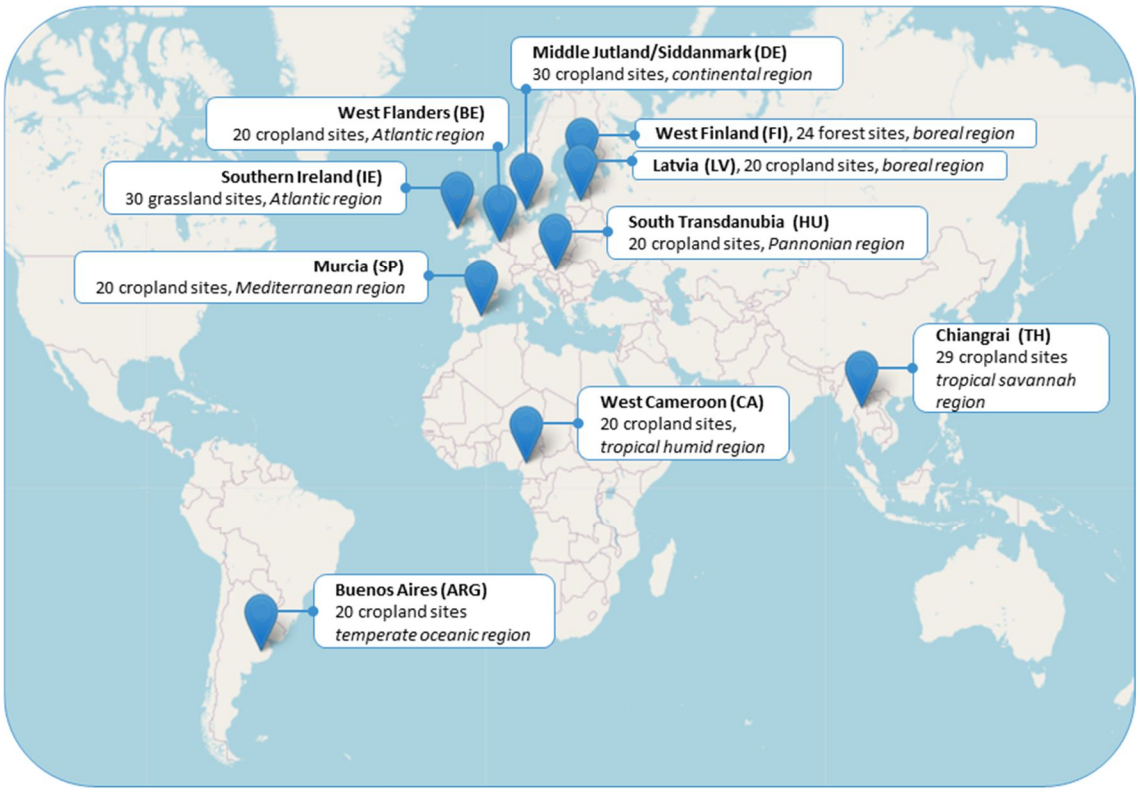


Figure 1. *Overview illustrating the geographical distribution of sampling sites of the cross-biome network of sites. The number of sites sampled in each region is included in each text box, along with the soil use type and, in italics, the name of the biogeographical region (Extracted from Deliverable 2.2.).*

At each region, we selected five plots for two soil management types: conventional and alternative (organic agriculture, grassland polycultures or continuous cover forestry depending on the biome), considering the sampling across two contrasting levels of soil degradation. To obtain our soil degradation index, we extracted the variables soil erosion and soil organic carbon (SOC) of each study site from available maps (full details in D2.1). We also obtained, from satellite information, characteristics of the landscapes surrounding our plots, namely the proportion of natural habitat nearby (%naturalness) and the number of hedgerows or treelines, as both have been previously shown to be important drivers of both agricultural biodiversity and functioning (full details for these measurements in ANNEX I and II). At each study site, we collected one composite sample out of five soil samples from the top 10 cm of the soil profile using soil cores of 5 cm of diameter. All soil samples were sieved (2-mm mesh) and air-dried for physico-chemical analyses (Table 3, section “Measurement of soil properties and functions”, below).

A soil subsample for each site was frozen at −20°C for DNA extraction and sequencing, phospholipid-derived and neutral-derived fatty acids (PLFA and NLFA), and functional nitrogen(N)-genes analyses (full details in D2.2, Soil biodiversity status in European and International biogeographical regions). Briefly, we determined the abundance of different soil taxonomic groups through PLFAs (i.e. bacteria, fungi, actinobacteria, methanotrophs, protists, microeukaryots, and total microbial biomass) and NLFAs (i.e. arbuscular mycorrhizal fungi, and total microbial storage biomass). The number of nematodes and micro-arthropods (collembola and acari) were also extracted from 100 g fresh soil using the Oostenbrink elutriator and cotton wool filter method. Abundance was calculated as the average of duplo counts. We also determined the abundance of root-feeding nematodes, as an additional indicator of an important function for croplands (root herbivory). Finally, soil DNA was extracted from freeze-dried soils using 0.25 gr of soil and the DNeasy *PowerSoil* HTP 96 Kit (Qiagen Inc., Valencia, CA, USA). “Species” (ASV) richness for eukaryotes (18S), prokaryotes (16S) and fungi (ITS r-RNA) were quantified through DNA metabarcoding with taxonomy markers for prokaryotes, fungi and non-fungal eukaryotes. We also compared two different approaches for soil biodiversity assessment through DNA surveys. The first approach was harmonized with the SOILBON European soil initiative (https://www.globalsoilbiodiversity.org/soilbon), targeting three ribosomal markers (1 pair of primers for

each one) to assess bacteria, archaea, fungi, protists, nematodes, micro-arthopods and earthworms: 16S V3-V4 region for prokaryotes, 18S V4 region with universal primers for Eukaryotes and ITS2 region specific for fungi. The second approach seeked deeper resolution for each of the soil organisms within the category of non-fungal eukaryotes. To do this, we targeted a specific gene (and specific region) for each type of soil organism: protists (18S V4-V5 region), nematodes (18S V6-V8 region), arthropods (COI gene) and earthworms (mit 16S region) (Table S1 ANNEX III).

3. Impact of soil degradation and soil management on soil biodiversity

The effects of high, medium or low soil degradation and conventional vs alternative soil management strategies on soil biodiversity were quantified in the 7 EU NUTS-2 and 3 international regions. We addressed this objective using multiple layers of information and levels of detail in the analyses.

First, we evaluated how conventional and alternative soil management impacted different soil biodiversity groups (i.e. alpha-diversity, mesofaunal groups and PLFA/NLFA concentration) measured. We then conducted a more comprehensive set of analyses to disentangle: i) the relative importance of soil management in relation to other major environmental factors in driving soil biodiversity and composition, and ii) potential interactions between soil management and these other environmental drivers.

3.1. Relative impact of soil management on biodiversity

We first assessed the effects of soil degradation and soil management strategies on soil biodiversity across the 7 EU NUTS-2 and 3 international regions. To achieve this, we calculated the Relative Interaction Index (RII, Armas et al. 2004) as a standardized response ratio for each of our soil biodiversity metrics (alpha-diversity and mesofaunal groups). This analysis allows us to produce a straightforward and accessible preliminary assessment of the impact of conventional vs alternative soil management on soil biodiversity. RII offers better statistical properties than the most commonly used log response ratio, while still providing a clear insight into how multiple variables respond to a given factor (in our case, conventional vs alternative agriculture). The RII was calculated as:

$$RII = (Xc - Xa)/(Xc + Xa) \quad (1),$$

where X is the variable of interest and Xc and Xa are the values on conventional and on alternative soil management respectively. This index ranges from −1 to 1, with RII values <0 representing situations in which values for alpha-diversity and mesofaunal groups are greater under alternative management. On the contrary, RII values >0 will represent higher values of biodiversity under conventional management.

Overall, alternative soil management benefited eukaryotes, particularly fungi (Figure 2), which is consistent with several reviews and meta-analyses concluding that alternative farming has an overall positive effect on biodiversity (Barral et al., 2015; Gomiero et al., 2011; Hole et al., 2005; Tuck et al., 2014). Both prokaryotes and fungi are essential for degradation of complex substrates and release of nutrients for plant uptake (Paula et al., 2020), which have direct effects on the soil health in agrosystems. However, conventional agricultural practices, such as fungicide application and tillage could destroy soil structure and fungal hyphal extension, which would consequently affect fungal community functioning (Sofo et al., 2022).

In contrast with these positive or neutral effects on soil microbes, we found a trend towards conventional soil management favoring the richness of Acari and Collembola. Different mechanisms could explain the positive effect of these practices on soil mesofauna. Firstly, the use of fertilizers and herbicides could favor the colonization of Acari and Collembolla opportunistic species such as bacterivores and omnivores. Moreover, George et al. (2017) showed a negative relationship between some groups of Acari (e.g. Mesostigmata) with soil C, responding favorably to a lower C:N ratio. According to our results, soil organic carbon is favored by alternative soil management (see Figure 25 below), while conventional agrosystems have a higher contribution of N due to fertilizers, so the decrease in the C:N ratio could be favoring both groups of organisms. Finally, Chang et al. (2013) found that both the richness and abundance of soil Collembola species increased in fertilizer and herbicide application treatments in a

soybean crop in a marsh in China, likely due to the increase in root and shoot biomass in these treatments. They also found that fertilizer application enhanced microbial resources and their grazers (e.g. springtails) by increasing nutrient availability to plants and soil.

CROPLANDS. ALTERNATIVE VS CONVENTIONAL

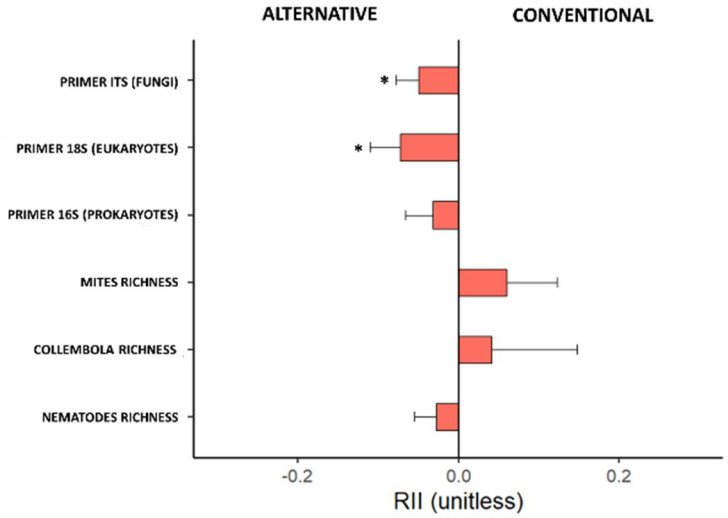
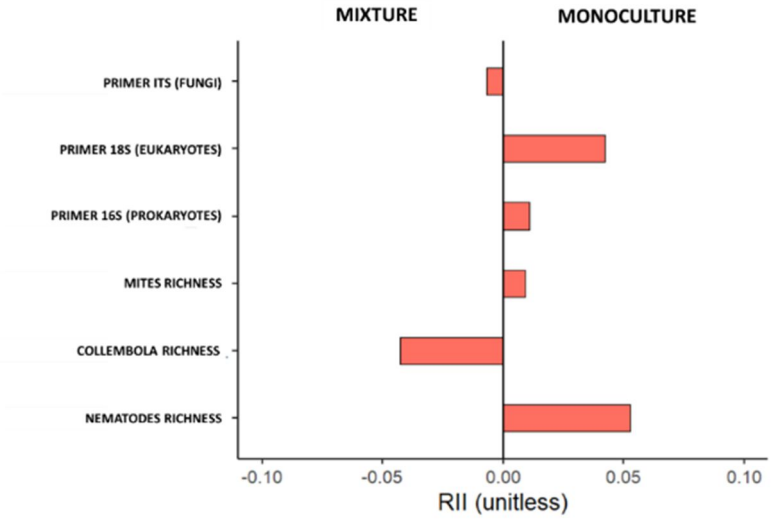


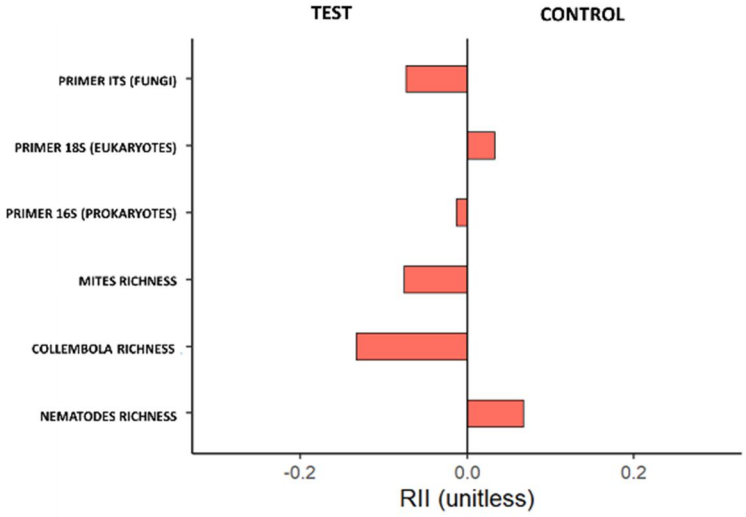
Figure 2. *Relative interaction index (RII) for the alpha-diversity (16S, 18S and ITS) and mesofaunal (i.e. richness of nematode, collembola and mites) measured in croplands for 5 EU NUTS-2 and 3 international regions (mean ± SE). RII was calculated for each of the 10 pairs of conventional vs alternative, irrespective of their soil degradation level. Negative RIIs show higher biodiversity levels in alternative vs conventional agriculture, whereas positive RIIs show the contrary. Asterisks indicate for which organisms these differences are significant (p value <0.05).*

Considering the differences in soil management between sites in Ireland and Finland, which involve grassland pastures (mixed plant species vs monoculture) and agroforestry (test vs control) in the case of Ireland, and forest systems with continuous vegetation cover or clear-cut forestry in Finland, independent analyses were conducted to determine the relative impact on each biodiversity group (Figure 3). Overall, there is consistency in the positive effect of alternative soil management in Finnish forest sites (i.e., continuous cover) for prokaryotes, eukaryotes, and fungi (16S, 18S, and ITS), but not for mesofaunal groups. In the case of Ireland, we found mixed results depending on the soil organisms, with greater diversity under the alternative soil management (polycultures or mixture) for springtails and fungi (ITS), but lower for eukaryotes and nematodes. It must be noted, however, that this last set of results are only trends, as all sites within a given management category were averaged to perform these comparisons, and therefore no error estimation or statistical tests were performed. Follow-up and specific detailed studies for these two biomes are currently being performed.

IRELAND. MIXTURE VEGETATION VS MONOCULTURE



IRELAND. TEST VS CONTROL (AGROFORESTRY)



FINLAND. CONTINUOUS COVER VS CLEAR CUT FORESTRY

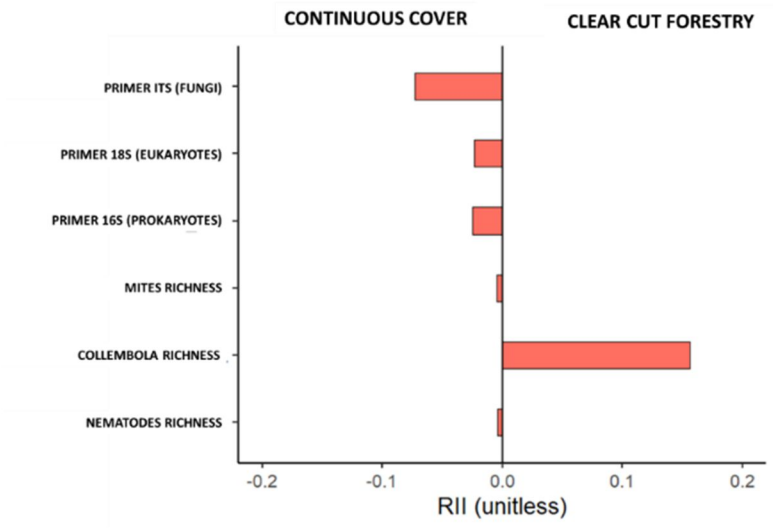


Figure 3. *Relative interaction index (RII) for the alpha-diversity (16S, 18S and ITS) and mesofauna (i.e. richness of nematode, collembolan and mites) measured in grasslands and agroforests of Ireland (above and middle) and forests of Finland (below). RII was calculated for each of the 10 pairs of mixture vegetation vs monoculture in grasslands and test vs control in agroforestry for Ireland, and continuous cover vs clear cut forestry for Finland. Negative RIIs show higher biodiversity levels in mixture and continuous cover agrosystems, whereas positive RIIs show higher biodiversity levels in monoculture and clear cut forestry.*

3.2. Multiple drivers of soil (alpha) biodiversity

Soil biodiversity not only depends on soil management practices, but also strongly responds to basic soil attributes such as its texture or pH, landscape context, climate and their interactions. To account for all these effects, we conducted linear models to determine the main predictors (soil, climate, and soil management) and their interactions on the taxonomic richness of the different soil groups. The most parsimonious model was selected using stepwise Akaike Information Criterion (stepAIC).

Permutation tests were used to avoid assumptions of normality. Hence, the initial model included the alpha-diversity of microbial (16S, 18S and ITS) and mesofaunal (richness of nematode, collembola and mites) groups, in addition to PLFA and NLFA concentration (bacteria, fungi, actinobacteria, methanotrophs, arbuscular mycorrhizal fungi, protists) as a response of: soil management x soil_degradation x aridity + %naturalness + surface_hedgerows + sand_content + soil_pH. Our response variables were the diversity and abundance of the different soil organisms measured in SOILGUARD, and an overall index of soil biodiversity (average of the standardized value for the richness of each group, as in Allan et al. 2014). All response variables were standardized by region, using Z-scores, to minimize potential confounding effects of unmeasured factors (e.g., different socio-economic or biogeographical backgrounds for each region). We repeated these analyses including only EU NUTS-2 cropland sites (alpha-diversity and mesofauna groups), and also including EU NUTS-2 and international cropland sites (only related to alpha-diversity due to lack of mesofaunal data in international regions). These analyses will be repeated and compared soon for the rest of the soil biodiversity groups from the second DNA strategy.

This approach allowed us to evaluate:

i) The interaction between soil management and soil degradation, after removing (as much as possible in an observational study) the influence of potential confounding factors such as different region-specific socio-ecological features or crop type,

ii) The dependence of the climatic context (interactions including aridity). This will provide complementary information to that of WP3, and also aid predictions on the future impacts of climate change on soil biodiversity,

iii) The influence of landscape attributes on soil biodiversity (%naturalness, surface hedgerows). These landscape features are also susceptible of being sustainably managed, are related to habitat-provisioning Nature's Contributions to People (NCPs; WP4) and could be promoted with agri-environmental schemes; hence we will produce results regarding their influence on soil biodiversity,

iv) The influence of soil type, other than soil management or degradation levels, on soil biodiversity (pH and texture).

For European NUTS-2 regions, the results show that the main predictor of biodiversity is aridity, as well as its interaction with soil management and soil degradation status (Table S1 in Annex IV, Figures 4 and 5). When analyzing the different groups of soil biodiversity, we observed that management is one of the main predictors for eukaryotes, collembola richness, and bacteria (PLFA concentration). With the exception of collembolans, alternative soil management had a positive influence on all mentioned groups after accounting for the influence of other well-known predictors of soil biodiversity. Moreover, fungi, arbuscular mycorrhizae, and actinobacteria were mainly determined by soil conditions such as texture, or pH. The influence of these soil conditions on microorganisms is well-documented. For instance, Cho et al. (2016) showed that soil pH significantly affects the community composition of these microorganisms, acting as a primary driver for their distribution and activity. Additionally, Xia et al. (2020) demonstrated that soil texture impacts microbial communities by influencing soil aeration, water retention, and nutrient availability, which are critical for the growth and function of microbes.

Collembola are a group of organisms that are usually present in high abundance in less disturbed sites (in particular, they are negatively affected in agricultural and horticultural sites) (Rutgers et al., 2009; Siepel, 1996; Siepel et al., 2018). However, some studies found that Collembola abundances were not affected by disturbance such as tillage and land use type (Martins da Silva et al., 2016; Parisi et al., 2005; Reeleder et al., 2006), and that their density and richness were even increased by fertilization and herbicide addition. These results could be explained by higher shoot and roots biomass, microbial resources and their grazers (i.e. the Collembola) in these intensive conventional treatments. Other studies found that more intensive management promoted the population of specific taxa (Benckiser, 1997). Collembola is a group of organisms that includes various taxa and functional groups, possessing diverse functional traits and adaptation strategies, that react in different ways to environmental and anthropogenic disturbances. As such, specific taxonomic or functional groups may provide more information or insight into the effects of management and land use than overall richness and/or total abundance of Collembola (George et al., 2017; Martins da Silva et al., 2016).These results are also consistent when considering the regions outside of the European Union (Argentina, Thailand, and Cameroon) (Figures 4 and 5; Table S2 in Annex IV). In this case, soil management was one of the main predictors for eukaryotes and methanotrophs. Conversely, if we consider average biodiversity (biodiversity index), the variation observed in croplands across the cross-biome network of sites is largely explained by soil degradation status and its interaction with soil management. The management x degradation interaction pointed towards more benefitial effects for soil biodiversity of alternative agriculture in soils that are already degraded.

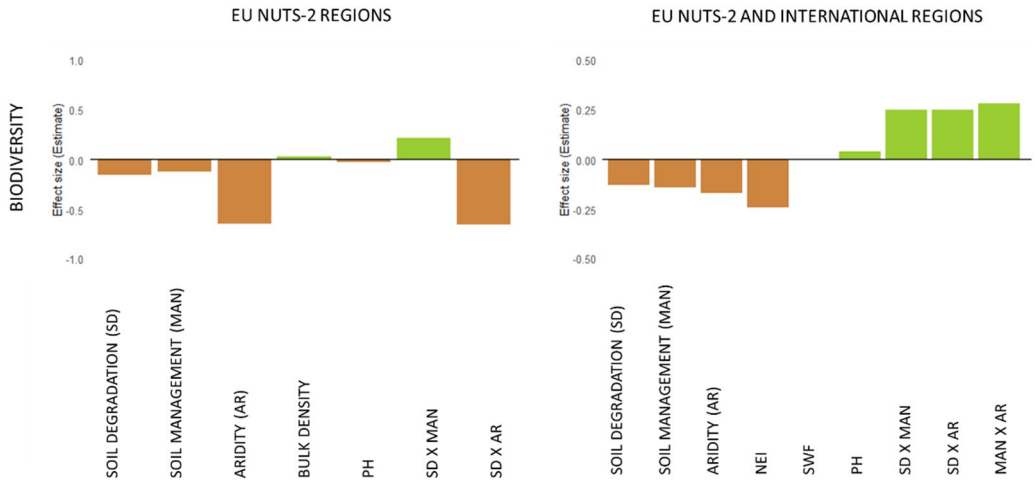


Figure 4. *Effect size of the different soil physico-chemical properties, landscapes attributes, soil agricultural management and degradation state and their interactions on the biodiversity (i.e. biodiversity index) for European NUTS-2 (left side) and European NUTS-2 and international regions (right side). Effect sizes (t-values) of the linear models are represented in green and brown for positive and negative effects, respectively. NEI: naturalness evaluation index; SWF: small wood features.*

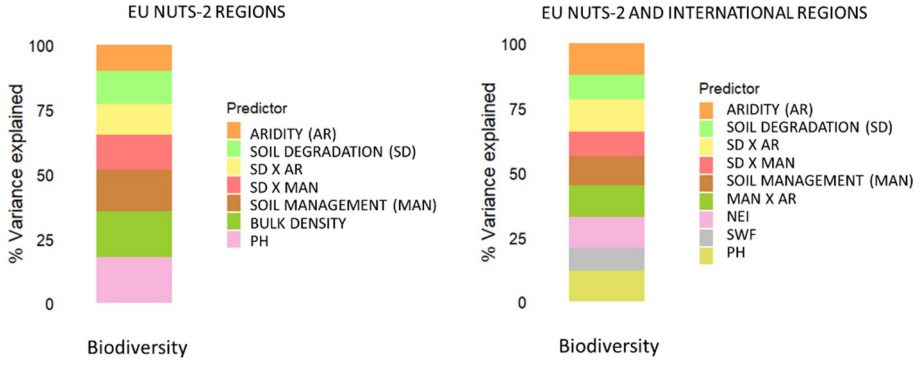


Figure 5. *Variance partitioning illustrating the relative importance of different soil physico-chemical properties, landscapes attributes, soil agricultural management and degradation state and their interactions on biodiversity (i.e. biodiversity index) for European NUTS-2 (left side) and European NUTS-2 and international regions (right side). NEI: naturalness evaluation index; SWF: small wood features.*

3.3. Factors driving changes in soil biodiversity community composition (beta-diversity and species-level analyses)

While the analyses above focused on the impacts of soil management, degradation and properties on local richness (alpha-diversity), they provide limited insight into which species benefit more under contrasting conditions. To assess these compositional changes in soil organisms in response to contrasting management, degradation and environmental conditions, we analyzed their Bray-Curtis index (beta-diversity). We calculated a Bray-Curtis dissimilarity matrix among samples with the vegan R package, using the 16S, ITS, 18SV4, 18SV6V8 and 16S mitochondrial amplicon sequencing data for prokaryotes, fungi, protist, nematodes/micro-arhtropods (collembola + mites) and annelids, respectively. The way these amplicon datasets were generated and the reasons why they were chosen for further analysis are explained in Annex III). We used Bray-Curtis dissimilarity as it handles the sparse sequencing datasets with many zero counts, and also to harmonize our results with those from LUCAS (Labouyrie et al 2023, Köninger et al. 2023). In any case, results were similar when using alternative dissimilarity distances (e.g.,

Jaccard index, both binary and no binary). We also considered other beta diversity indexes based on phylogenetic distance (such as weighted and unweighted UniFrac), but the dimensionality of our data (six amplicon datasets) did not allow us to proceed further. This further exploration will be considered for a specific analysis of beta diversity (including its nestedness and turnover compartments), which is ongoing and planned to be part of a SOILGUARD publication during this last year of the project.

Prior to the calculations, we divided the datasets by region, except for the EU croplands ones (West Flanders - BE, Latvia – LV, Murcia – ES, Middle Jutland/Syddanmark – DE, South Transdanubia - HU), which were analysed together to identify general patterns. To evaluate the effects of country (only in the case of EU croplands), soil management, and land degradation on the community composition of microbes (prokaryotes, fungi), microfauna (protists, nematodes) and mesofauna (micro-arthropods: acari, collembola and annelids), we first performed a two or three-way PERMANOVA with the Adonis R package (999 permutations, $p < 0.05$). To reduce the dimensions of the dissimilarity matrix and facilitate pattern observation, we performed a constrained analysis using the combination of management x degradation as the constraining factor. This was done by applying a canonical analysis of principal coordinates (CAP) (Anderson and Willis, 2003) with the CAPdiscrim function in the R package BiodiversityR (Kindt and Coe, 2005). Additionally, we assessed the abundance of specific species (ASVs) that were significantly associated with each contraining factor, thereby helping to explain some of the differences between conditions. These ASVs were considered significantly correlated with CAP ordination scores by using the envfit function from the vegan package V2.6-4 999 permutations, when p value (after false discovery rate correction) < 0.1 and r2 > 0.3.

3.3.1. EU cropland soils: West Flanders (BE), Middle Jutland/South Denmark (DE), South Transdanubia (HU), Latvia (LV), Murcia (ES)

PERMANOVA on beta-diversity shows that the factor which impacts soil biodiversity community composition the most (7 types of organisms) is the geographical location ($p < 0.05$; explained variability 10-27%). In terms of soil prokaryotes community composition, soil degradation and soil management effect are EU NUTS-2 region dependent ($p < 0.05$; 2-4% variability), while no significant effects are observed on soil prokaryotes when management and degradation are combined ($p > 0.05$; 2% variability). In contrast, fungi, protist, and nematode community compositions are significantly affected by soil management ($p< 0.05$; 1-2% variability), being dependent on degradation only when considering the EU NUTS-2 region factor (EU NUTS-2 x management x degradation: $p < 0.05$; 2% variability), except for protists. Soil management and soil degradation status do not significantly impact the community composition of micro-arthropods. Although this may seem contradictory, these results may tell us two things: DNA techniques for mesofauna are not yet strong enough to track their community composition well (see section 3.5.1 below), or that other management strategies other than the conventional vs alternative dichothomy considered in our analyses (e.g., no-tilling) are modulating the response of these organisms (Betancur-Corredor et al., 2022). Annelids were discarded for all subsequent analyses due to the low number of successfully sequenced samples (n < 50%).

A more in-depth analysis of the relations between soil management, soil degradation and soil biodiversity has been performed through assessing the ASVs that fit into the ordination plot after CAP analysis. Our results show that four ASVs from the prokaryotic and one from the fungal datasets were correlated significantly with sample disposition of beta-diversity constrained analysis (Figure 6). Those belonging to the orders Azospirillales (ASV413, uncultured family), KD4-96 (ASV786, uncultured family), **Gaiellales** (ASV991, uncultured family), and Sebacinales (ASV110, genus: ***Stachybotrys***) are **related to conventional management samples under low degradation levels**. Overall, the differential abundance of these ASVs, when compared to CAP analysis, can indicate that these microbes are responsible for the differences observed between the various levels of soil degradation. The Gaiellales order and the genus *Stachybotrys* are correlated with consecutive fertilization treatments, and consequently higher available N content in soils (Ning et al., 2021; Wang et al., 2024), which is in line with the higher N content observed in the conventional fields in our study. In addition, the ASV1529 belonging to **Frankiales (family:**

Geodermatophilaceae) fits into samples with low degradation levels but under alternative management.

.

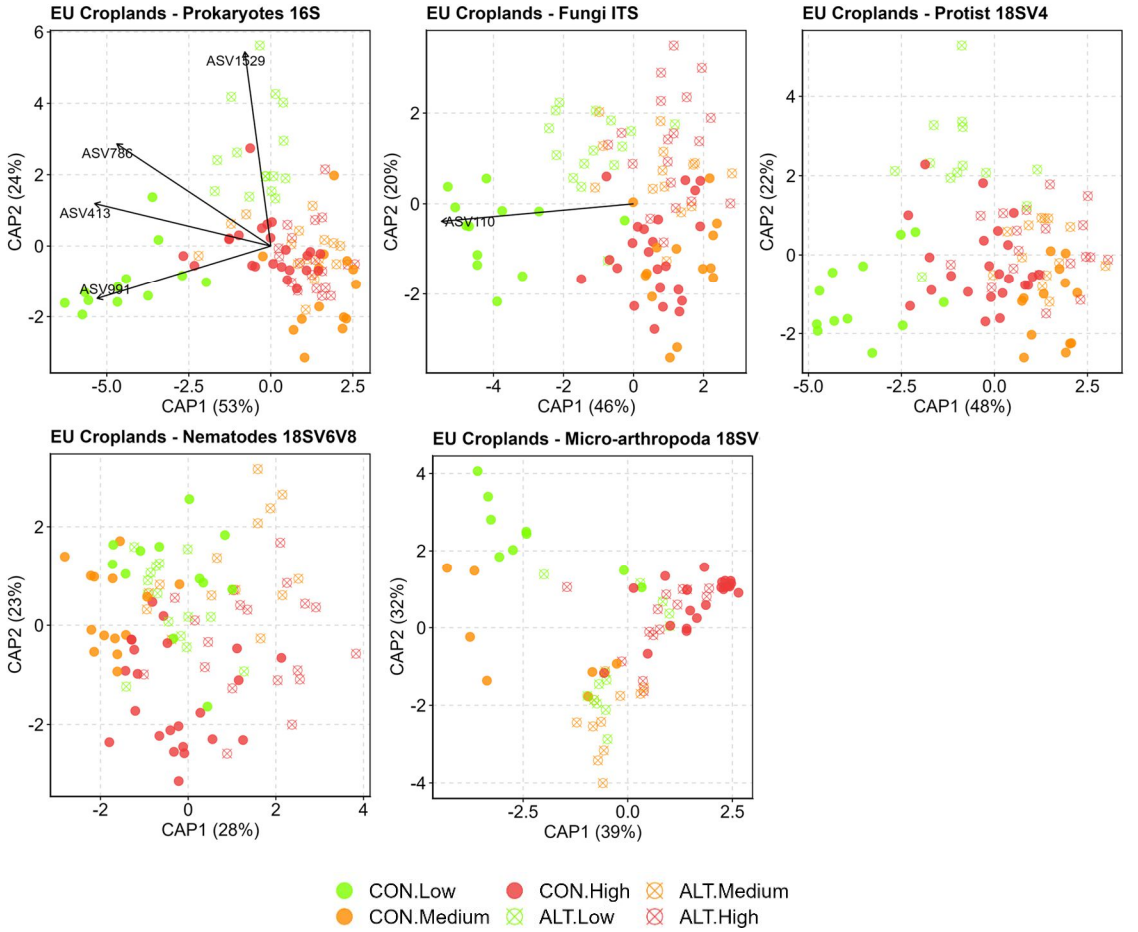


Figure 6. *Beta-diversity EU croplands analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been colored and shaped by the combination of Management and Degradation. Different sample size was obtained depend on the target taxa → Prokaryotes (16S): CON.Low=12, CON.Medium=14, CON.High = 24, ALT.low=15, ALT.Medium=15, ALT.High=20; Fungi (ITS): CON.Low=12, CON.Medium=14, CON.High = 24, ALT.low=15, ALT.Medium=14, ALT.High=20; Protists (18SV4): CON.Low=12, CON.Medium=13, CON.High = 24, ALT.low=13, ALT.Medium=13, ALT.High=20; Nematodes (18SV6V8): CON.Low=12, CON.Medium=14, CON.High = 24, ALT.low=15, ALT.Medium=14, ALT.High=20; Annelida (16Smit): these were discarded due to number samples < 50% of the total (n=44). Micro-arthropods (18SV6V8): CON.Low=9, CON.Medium=7, CON.High = 23, ALT.low=13, ALT.Medium=11, ALT.High=15. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

3.3.2. International croplands: Chiangrai (TH), Buenos Aires (ARG), West Cameroon (CM)

Moving to the international cropland regions, impacts at beta-diversity level by management/degradation were similar to those found for EU croplands. First, **Chiangrai's** soil prokaryotes beta-diversity is significantly impacted by the level of soil degradation (variability explained 13%), whereas global management (variability 5%) or the combination of management and degradation (variability 9%) do not significantly impact the composition ($p > 0.05$).

Moreover, fungal, protistan and nematode beta-diversities were affected by the interaction between degradation level and management (management x degradation $p < 0.05$; 9-10% variability explained). Samples belonging to alternative vs conventional management are more distant when the

degradation level is low/medium across the main axis of the CAP1, than under high degradation levels. In contrast, management and degradation did not significantly affect the community composition of micro-arthropods.

Assessing the ASVs that fit into the ordination plot after CAP analysis, we found 53 significant correlations from the prokaryote dataset (only the top 10 represented), three from the eukaryote datasets when computed altogether, and one from micro-arthropods (Figure 7). In terms of the prokaryotes, one ASV from the order Anaerolineales (ASV2016), fit into the moderatelly degraded soils under alternative management. Otherwise, ASVs from Rhizobiales (ASV2283/1145; family: Xanthobacteraceae), KD4-96 (ASV341; family: KD4-96), Isosphaerales (ASV2106/3003/4069/1261; family: Isosphaeraceae), Micrococcales (ASV1430; genus: *Sinomonas*), and Corynebacteriales (ASV3452; genus: *Mycobacterium*) were correlated with the disposition of samples belonging to low and high degraded soils under alternative management. Regarding the micro-arthropods, the ASV936 from Sarcoptiformes (unclassified family) fit into the low degraded soils under alternative management samples composition.

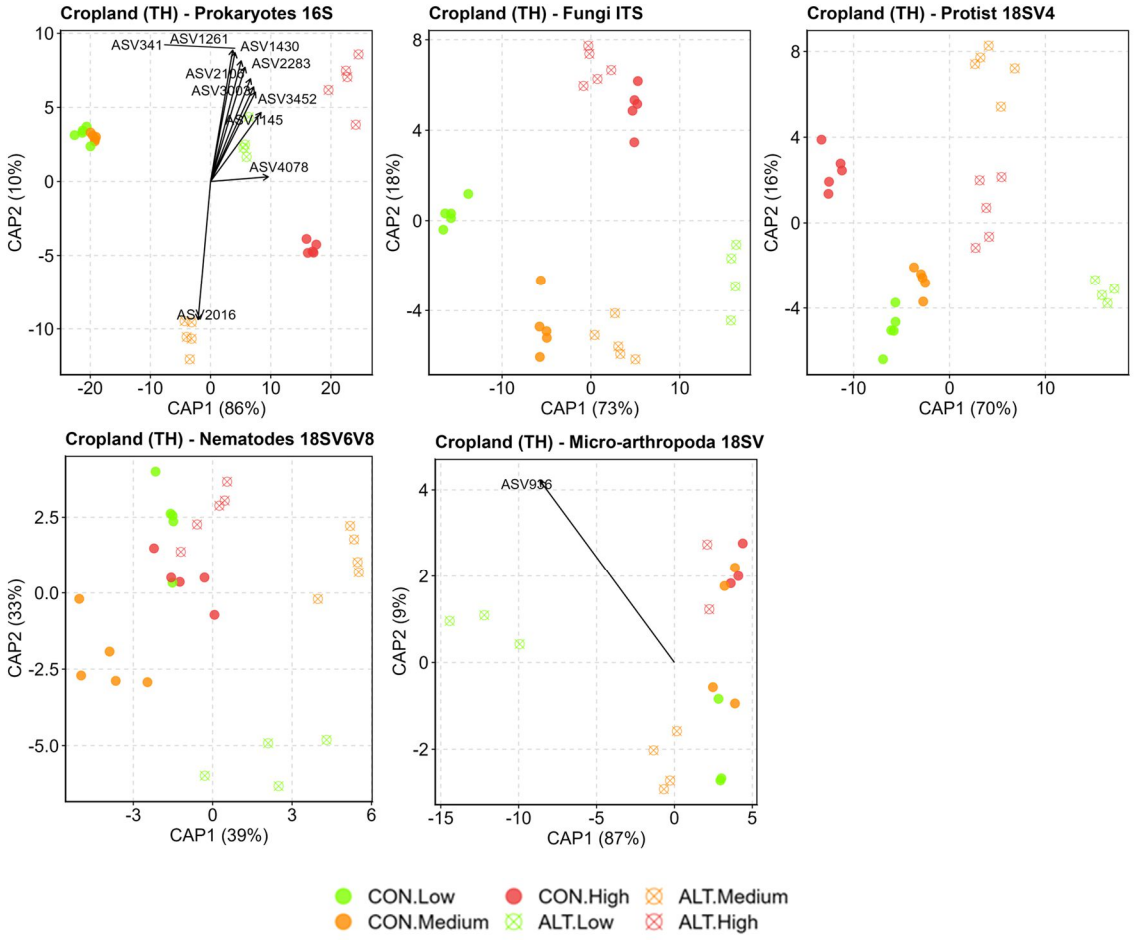


Figure 7. *Beta-diversity Chiangrai (TH) soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been colored and shaped by the combination of Management and Degradation.* Different sample size was obtained depend on the target taxa → Prokaryotes (16S): *CON.Low=5, CON.Medium=5, CON.High = 5, ALT.low=4, ALT.Medium=5, ALT.High*=5; Fungi (ITS): *CON.Low=5, CON.Medium=5, CON.High = 5, ALT.low=4, ALT.Medium=5, ALT.High*=5; Protists (18SV4): *CON.Low=5, CON.Medium=5, CON.High = 5, ALT.low=4, ALT.Medium=5, ALT.High*=5; Nematodes (18SV6V8): *CON.Low=5, CON.Medium=5, CON.High = 5, ALT.low=4, ALT.Medium=5, ALT.High*=5; Annelida (16Smit): these were discarded due to number of samples < 50% of the total (n=8). Micro-arthropods (18SV6V8): *CON.Low*=3, *CON.Medium=4, CON.High = 3, ALT.low*=3, *ALT.Medium=4, ALT.High=2*. *Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

For the degradation levels of Buenos Aires and Cameroon, we grouped the samples under moderatelly and highly degraded soils to have a balanced dataset in these regions. Similarly to EU

croplands and Chiangrai, prokaryote beta-diversity in **Buenos Aires** is significantly impacted by the level of soil degradation ($p < 0.05$; variability 9%), but not by management or the interaction between management x degradation ($p > 0.05$; variability 4%).

Management significantly affected the beta-diversity of fungi ($p < 0.05$; explaining 6-7%), independently of degradation level. Protist betw-diversity was significantly affected by both management ($p< 0.05$; variability 6%) and degradation ($p < 0.05$; variability 7%) as independent factors. Neither management nor degradation significantly affected the beta-diversity of nematodes and micro-arthropods.

Assessing the ASVs that fit into the ordination plot after CAP analysis, we found 19 significant from prokaryotes and 3 significant ones from micro-arthropods dataset (Figure 8). From Prokaryotes, several ASVs from Bacillales (ASV35784/98; genus: Lysinibacillus/Unclassified), Rubrobacterales (ASV1010/4216) and Actinobacteria phylum (ASV1187; class: unclassified) are related to high degraded soils disposition. In the other hand, Tepidisphaerales (ASV305; genus:WD2101_soils group), Corynebacteriales (ASV752, genus: *Mycobacterium)* and Gaiellalles (ASV3614; genus: *Gaiella*) are related to low degraded soils. Micro-arhthropods ones fit into low degraded soils under conventional management and belonged to Sarcoptiformes order (ASV488/1235/7945; family: unclassified).

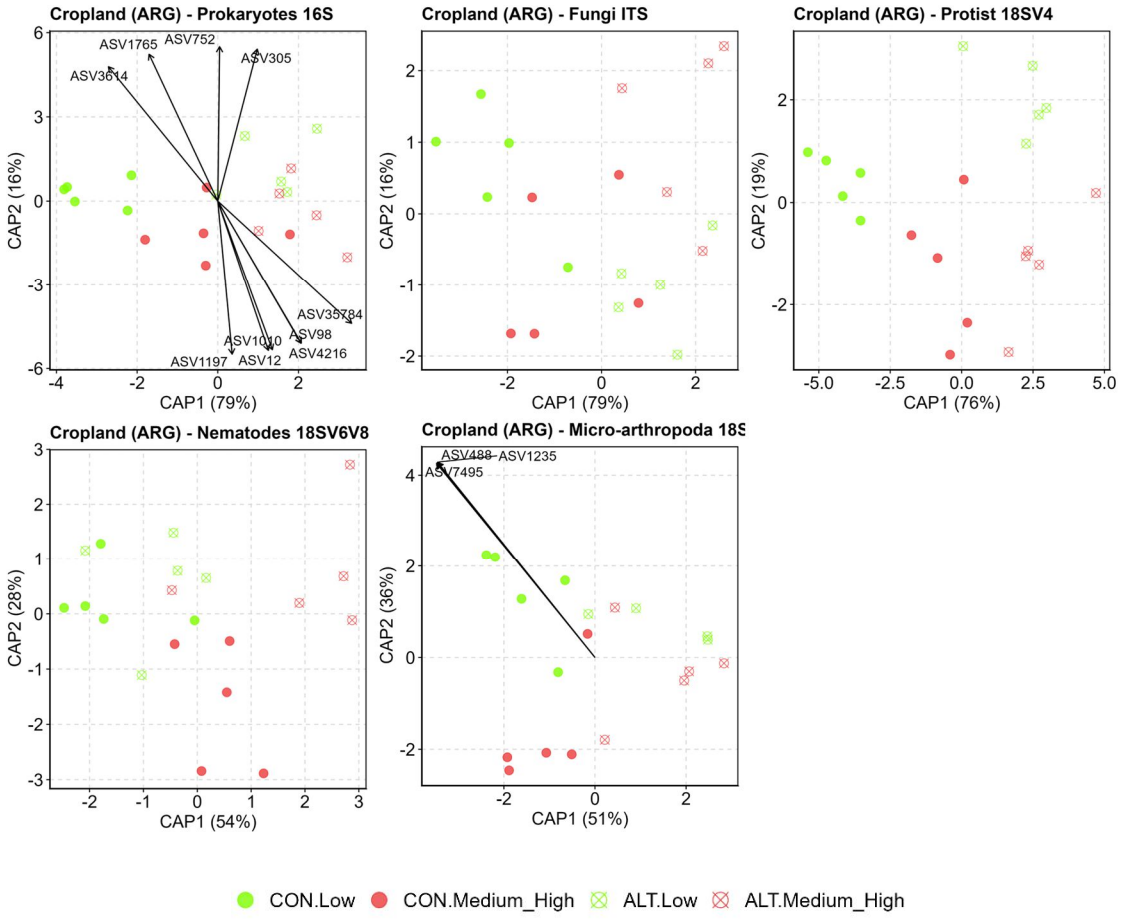


Figure 8. *Beta-diversity Buenos Aires soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been colored and shaped by the combination of Management and Degradation.* Different sample size was obtained depend on the target taxa → Prokaryotes (16S): *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Fungi (ITS): *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Protists (18SV4): *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Nematodes (18SV6V8): *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Annelida (16Smit): these were discarded due to number of samples < 50% of the total (n=6). Micro-arthropods (18SV6V8): *CON.Low=5, CON.Medium_high=5, ALT.low=4, ALT.Medium_high=5. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with*

LDA scores were represented by arrows. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.

Conversely to the previous cropland regions, the beta-diversity of prokaryoted in **Cameroon** was significantly impacted by management ($p < 0.05$; variability 10%), independently of land degradation level ($p > 0.05$; variability 4%). Similar results were observed in fungal community composition. The rest of eukaryotes were not affected by either management or degradation ($p > 0.05$; variability 4-22%).

Assessing the ASVs that fit into the ordination plot after CAP analysis, we found 29 significant correlated from prokaryotic datasets (only the top 10 represented) (Figure 9). Considering the top 10 most correlated, ASVs from Rhizobiales (ASV198/63; family Xanthobacteraceae) and KD4-96 (ASV5727; family: KD4-96) fits into the ordination of alternative management under both degradation levels. ASVs from AD3 (ASV921; family: AD3), Ktedonobacterales (ASV397/575/2310; genera: JG30a-KF-32/FCPS473), Sphingomonadales (ASV1587; family: Sphingomonadaceae), and Frankiales (ASV101, family: Geodermatophilaceae) fit into low degraded soils under conventional management. The signifcative correlation of these ASVs can be related with the significant differences observed in prokaryotes beta-diversity between managements. Which allow us to suggest that members from Ktedonobacterales, Sphingongomonadales, Rhizobiales and KD4-96 are more sensitive to different managements in West Camerooon.

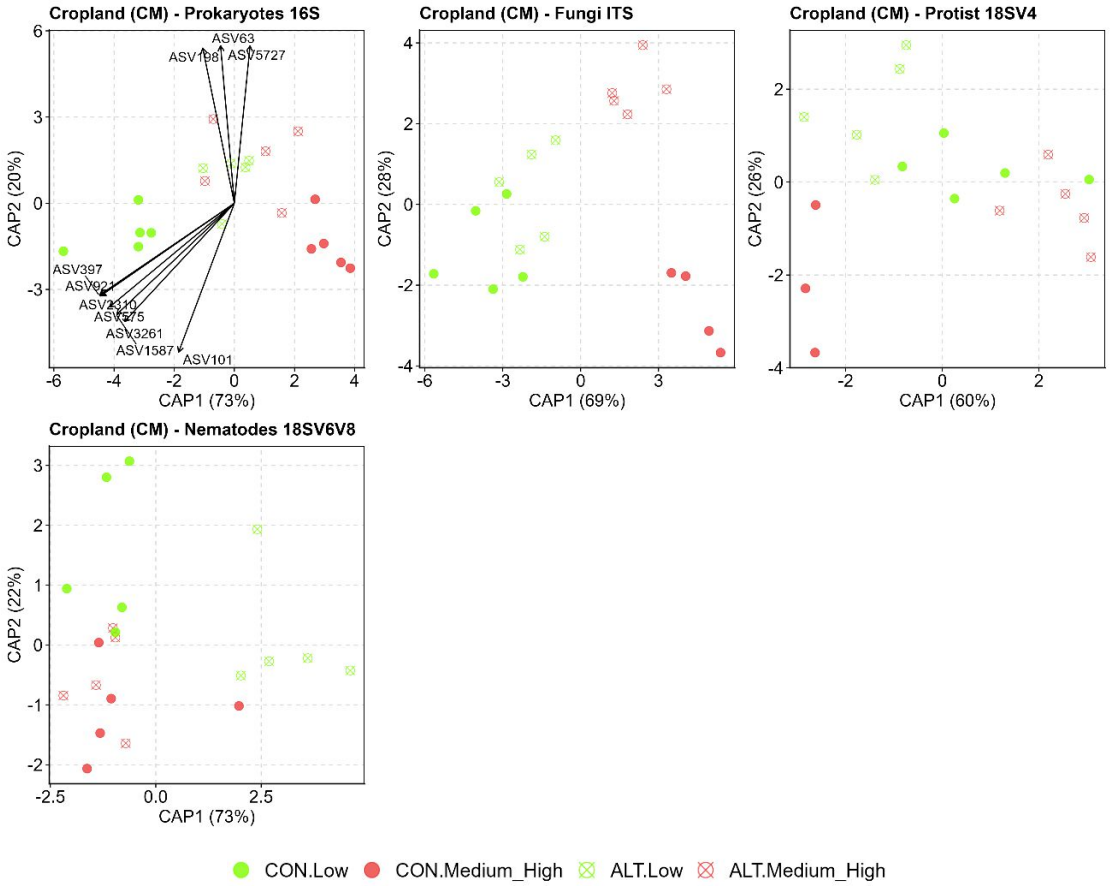


Figure 9. *Beta-diversity Cameroon soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been colored and shaped by the combination of Management and Degradation. Samples have been colored and shaped by the combination of Management and Degradation.* Different sample size was obtained depend on the target taxa → Prokaryotes: *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Fungi: *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Protists: *CON.Low=5, CON.Medium_High=5, ALT.low=4, ALT.Medium_high=5*; Nematodes: *CON.Low=5, CON.Medium_High=5,*

ALT.low=4, ALT.Medium_high=5; Annelida: these were discarded due to number of samples < 50% of the total (n=4). Micro-arthropods: these were discarded due to number of samples < 50% of the total (n=9). *Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

Overall, the effects of management and soil degradation in the croplands located in the three international regions were dependent on the type of soil organism and the specific region. The beta-diversity of prokaryotes was only affected by management in West Cameroon. Soil prokaryotic community composition was affected by soil degradation level in all the regions studied. These results may indicate a major sensitivity of prokaryotic community composition to secondary effects of soil degradation, such as soil acidification, a well-known factor influencing prokaryotic communities (Rousk et al., 2010). In contrast, the lack of degradation impacts observed in West Cameroon soils could potentially indicate a greater buffering capacity of their soils against these effects (all soils have pH between 5-6). Further, the consistent ploughing application across all sites in this region may reduce the noise in the system and allow for better observation of the impacts of organic fertilization on soil prokaryotic communities. However, the higher altitude of sites in Cameroon compared to the sites in other countries (1300-1600 m) could also affect the dynamics of prokaryotes. These differences in composition between degradation levels or management were linked to different ASVs depending on the region, which mostly belonged to Rhizobiales (*Bradyrhizobium*, *Pedomicrobium*, uncultured) Corynebacteriales (*Mycobacterium)*, Gailealles (*Gaiella*, unclassified) and Frankiales (*Acidothermus*, *Jatrophihabitans* and *Geodermatophilus*, unclassified), all from Proteobacteria and Actinobacteriota phyla with different potential roles in soil systems (Orgiazzi et al., 2016).

Regarding the eukaryotes, soil fungal community composition was mostly affected by soil management, though also significantly affected by soil degradation in the EU croplands and Chiangrai region. Protist community composition was affected by management and soil degradation in all regions except for West Cameroon. The sensitivities of both eukaryotic groups have been previously described in a detailed study about N fertilization and tillage intensities impacts in macro- and micro-aggregates (Pellegrino et al., 2021). Moreover, the intrinsic interplay observed by Pellegrino et al. (2021) between these groups, the soil structure, and C cycling calls for further study under this project, combining the co-occurrence network approaches and links between biodiversity and multifunctionality of croplands sites. Nematodes were sensitive to soil management in EU croplands and Chiangrai, with no ASV significantly associated to these changes. Although these three groups of eukaryotes were significantly affected by management, no ASVs were identified as the main drivers of these differences. This may be due to the variability between samples in terms of ASV abundance within the same condition, leading to weak correlations. Annelids and micro-arthropods were not affected by soil management or degradation, but only by geographic location (region).

These results further support the idea that eukaryote biodiversity, in particular fungi, is the most responsive group to soil management, being also significantly impacted by soil degradation level in some regions. Moreover, we emphasize that the current methods used to monitor annelids/earthworms or micro-arthropods are not optimal for cropland samples, as several samples were lost due to the low presence of reads belonging to these groups. In view of these results, discussions should arise about the way of soil sampling for these groups, hypothesizing that selecting identified transit zones for annelids and earthworms could be key to improving DNA-based ecological surveys.

3.3.3. Grasslands/agroforestry: Southern Ireland (IE)

Prokaryotic beta-diversity in the grasslands of Southern Ireland was significantly impacted by the level of soil degradation ($p < 0.05$; variability 10%), whereas soil management type ($p > 0.05$; variability 3%) did not significantly impact their community composition. This was also true for fungi, nematodes and annelida beta-diversity [soil degradation level ($p < 0.05$; explaining 10-16% variability)]. Protist and micro-arthropod communities were not affected by degradation ($p > 0.05$; 4-6% variability) or management ($p > 0.05$; variability 4-7%).

Assessing the ASVs that correlate with the ordination axis after CAP analysis, we found 10 significant correlations from prokaryotic datasets (Figure 10). Two of these ASVs were significantly related to the beta-diversity found in under low soil degradation levels, which belonged to Chthoniobacterales (ASV725; genus: *Candidatus_udaeobacter*) and Corynebacteriales (ASV563, genus: *Mycobacterium*) orders. Otherwise, ASVs from Rizhobiales (ASV20, family: Xanthobacteraceae), KD4-96 (ASV68, family: KD4-96), Propionibacteriales (ASV107, genus: *Nocardioides*), Gaiellales (ASV125; family: uncultured), Gitt-GS-136 (ASV170, family: Gitt-GS-136) and IMCC26256 (ASV338, family: IMCC26256) are significantly related to the beta-diversity found in sites belonging to moderate soil degradation levels.

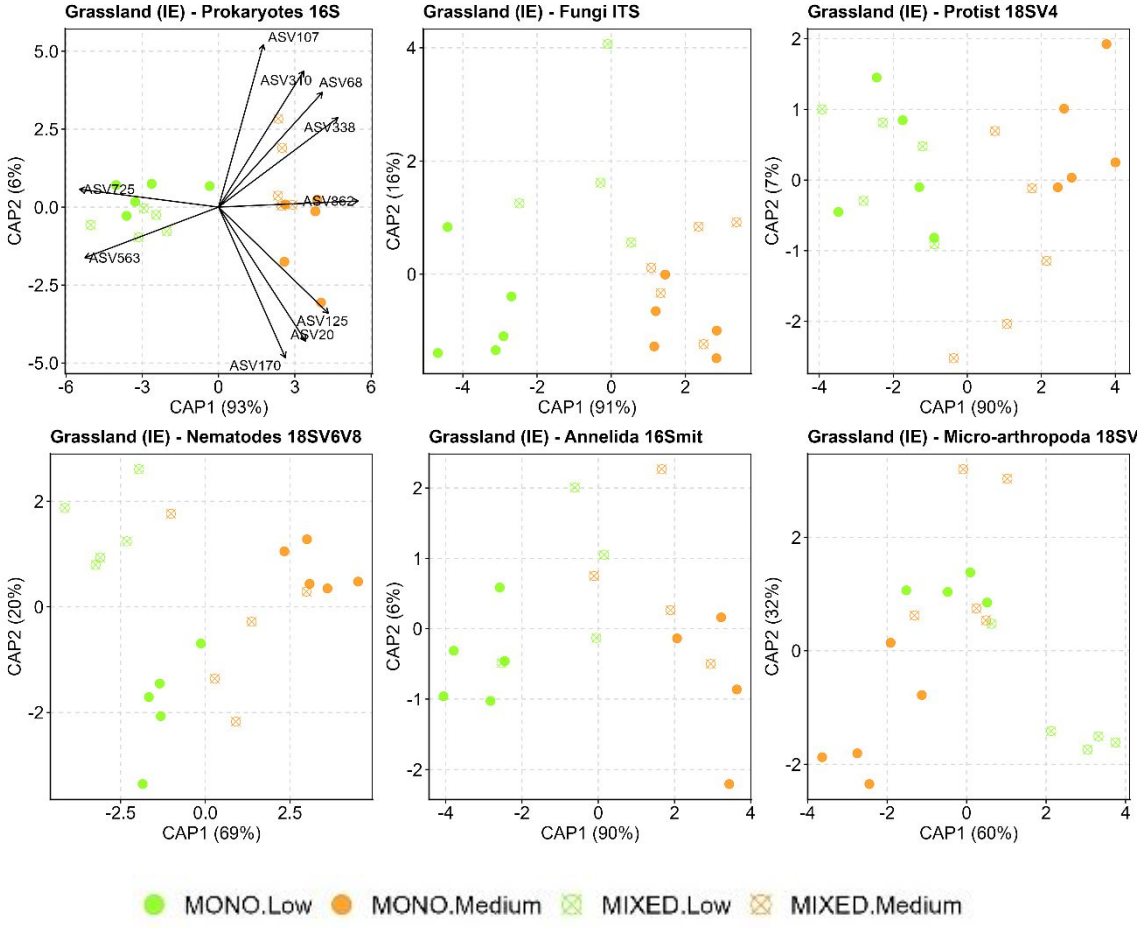


Figure 10. *Beta-diversity Southern Ireland grassland soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been coloured and shaped by the combination of Management and Degradation. Different sample size was obtained depend on the target taxa → Prokaryotes: MONO.Low=5, MONO.Medium=5, MIXED.low=5, MIXED.Medium=5; Fungi: MONO.Low=5, MONO.Medium=5, MIXED.low=4, MIXED.Medium=5; Protists: MONO.Low=5, MONO.Medium=5, MIXED.Low=5, MIXED.Medium=5; Nematodes: MONO.Low=5, MONO.Medium=5, MIXED.low=5, MIXED.Medium =5; Annelida: MONO.Low=5, MONO.Medium=4, MIXED.low=4, MIXED.Medium=5. Micro-arthropods: MONO.Low=4, MONO.Medium=5, MIXED.low=4, MIXED.Medium=5.Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

The beta-diversity of prokaryote and all eukaryote groups studied under the agroforestry sites in Southern Ireland did not show any significant impact from either management or soil degradation ($p > 0.05$; variability 6-21%). Albeit, we must consider the low sample size under this type of agroecosystem and the high dispersion observed in terms of composition between conditions (Figure 11), which can mask the identification of any effect.

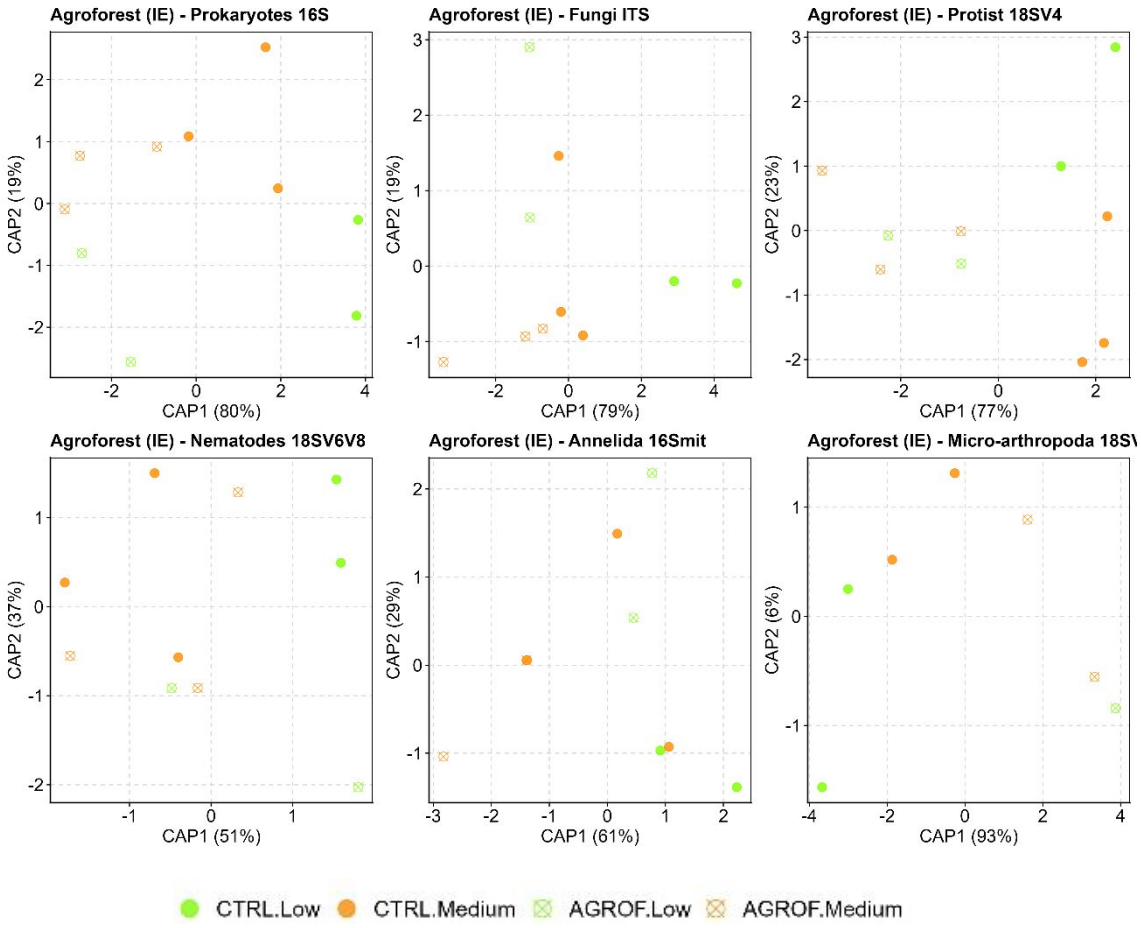


Figure 11. *Beta-diversity Southern Ireland agroforestry soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been coloured and shaped by the combination of Management and Degradation.* Different sample size was obtained depend on the target taxa → Prokaryotes: *CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium=3*; Fungi: *CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium=3*; Protists: *CTRL.Low=2, CTRL.Medium=3, AGROF.Low=2, AGROF.Medium=3*; Nematodes: *CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium =3*; Annelida: *CTRL.Low=2, CTRL.Medium=3, AGROF.low=2, AGROF.Medium=1*. Micro-arthropods: *CTRL.Low=2, CTRL.Medium=2, AGROF.low=1, AGROF.Medium=2. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

In contrast to the croplands studied, the effect of soil degradation in grasslands from Southern Ireland was consistent across all soil organisms except for micro-arhtropods. Conversely, agroforestry community composition was not significantly impacted by management nor degradation. While the agroforestry vs control comparison does not reveal clear patterns or statistical significances, combining these 5 sites with all grassland samples (mono/mixed, N = 25) shows a significant impact of agroforestry management on community composition, compared to the other two (MONO vs AGROF: $p < 0.05$, variability 8-16%; MONO vs MIXED: $p = 0.001$, variability 14-23%). However, micro-arthropods were different only between monoculture and agroforestry and annelids showed no significant differences between any pairwise comparison. No significant effects of soil degradation were observed in Irish samples altoghether (AGROF + MONO + MIXED). This indicates that while grassland sites in Ireland are strongly influenced by soil degradation, more contrasting managements such as agroforestry are the main driver of soil biodiversity changes, as previously shown for fungal communities (Bainard et al., 2012).

Significant correlations between ASVs and contrasting management x degradation levels were found only for prokaryotes. We found clear associations between two ASVs belonging to Corynebacterales (*Candidatus udaeobacter* and *Mycobacterium)* under low degraded soils, and Rizhobiales, Propionobacterlaes and Gaiellalles under soils with higher degradation (medium degraded soils). For

eukaryotes, correlations were significant but not strong enough (adjusted p.adjusted < 0.1) to be discussed here. This may be the cause of the small dataset (n = 20) and the higher variability in terms of eukaryotes´ ASV abundance between plots.

3.3.4. Forest: West Finland (FI)

For forests, we assessed the composition of each specific management within clear-cut conditions (FOLD: clearcut old, FMID: clearcut middle, FYM: clearcut mounding,) and continuous cover (CCR: continuous retained, CCGA: continuus gap, CCS: continuous single). Meanwhile different clear-cutting conditions refers to how long these have been applied, CCR refers to continuous maintenance of structural heterogeneity within a stand (species, tree age, and tree size), CCGA to continuos managements where gaps with a diameter of 10-60 m are established for regeneration (dense patches are thinned between gaps to maintain tree vigour and stability) and CCS to selectively harvest trees individually or in small groups over time, maintaining a continuous cover

The prokaryote and all eukaryotic group compositions in the forest soils of West Finland are strongly impacted by soil management ($p < 0.05$; variability 29-39%). When analyzing the correlated ASVs with CAP coordinates, we found 118 significant ASVs from prokaryotes (top10 represented), 27 from fungi (top10 represented), 11 from nematodes, 2 from annelids, and 10 from micro-arthropods (Figure 12)

Regarding the top 10 represented prokaryotes, two ASVs belonging to Rhizobiales (ASV288/1220; family: Xanthobacteraceae) relate to the FYM and CCGA managements. Meanwhile, ASVs from Corynebacteriales (ASV44265/809/567; genus: *Mycobacterium*), Isosphaerales (ASV80; family: Isosphaeraceae), Chitinophagales (ASV1590; genus: *Puia*), and Tepidisphaerales (ASV1126/978; genus: *WD2101_soil_group*) fit into FOLD and FMID managements. Regarding fungi, one ASV from Agaricales (ASV2660, family: Clavariaceae) fit into CCGA management. Otherwise, several ASVs from Leotiomycetes class (order unclassified), Helotiales (ASV2398/1312/1438/13710 family unclassified), and Sebacinales (ASV5313; family unclassified) were linked to FYM composition. Regarding the nematodes, the ASV306 from Dorylaimida order (unclassified family) and the ASV1663 unclassified at class level, are linked to CCS. Meanwhile Rhabditida (ASV2908/440/15823 genera: *Teratocephalus*/*Bunonema*) and Chomadorida orders (ASV950; genus: *Achromadora*), or unclassified ones (ASV17265/3046) are link to CCGA.

As for annelids, both ASVs belonged to the Enchytraeida order, the ASV1 from *Chamaedrilus* genus has no clear association to any sample distribution, and the ASV8 belonged to *Cognettia* follow an opposite trend to the CCS management samples. For micro-arthropods, the ASV101 from Mesostigmata order (family: Zerconidae) and ASV290 from Sarcoptiformes (family: Steganacaridae) fit into FOLD samples. Two ASVs from Sarcoptiformes order (ASV23/7443; genus: *Tectocepheus*) fit into CCR samples. Finally, several ASVs, also from Sarcoptiformes order (ASV96/1941/6660/12698/17048/17181; genus: *Nothrus*/unclassified), fit into the FYM management ones.

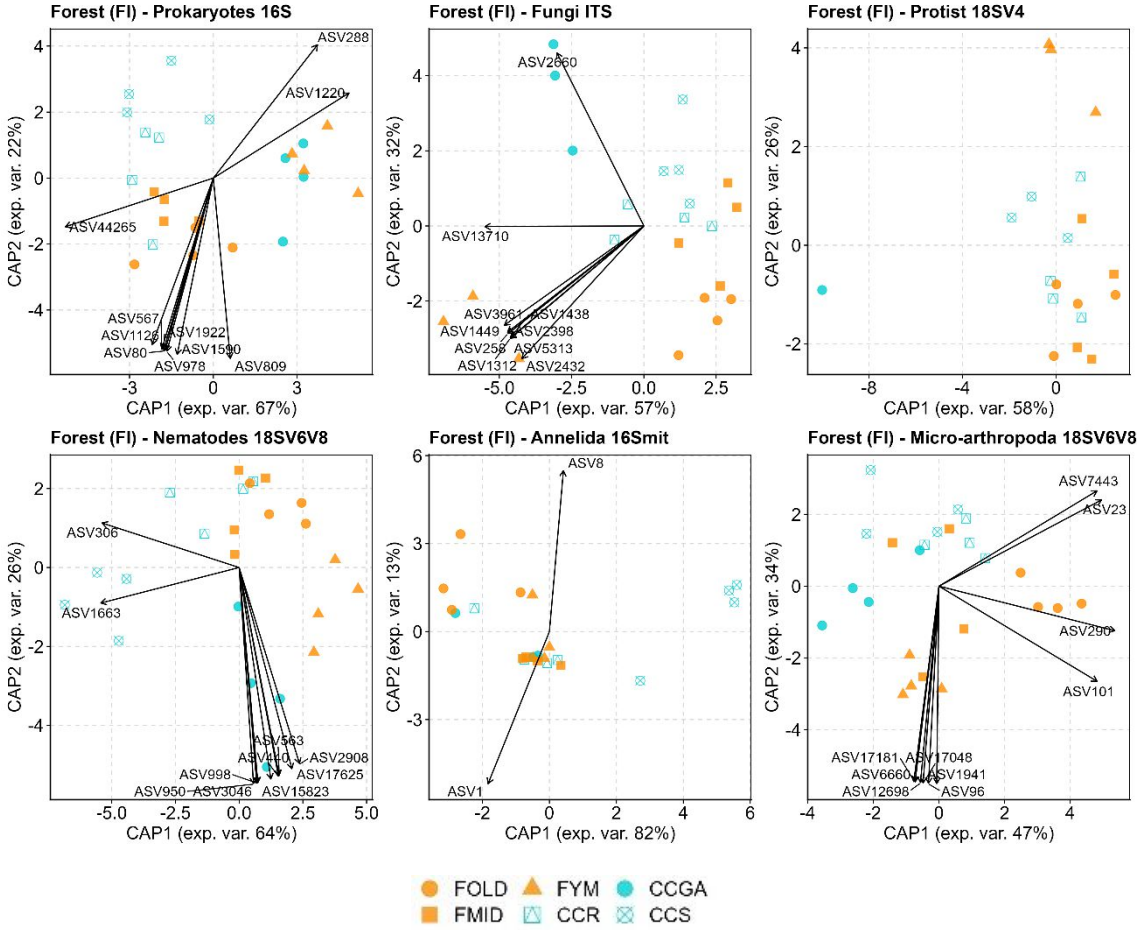


Figure 12. *Beta-diversity West Finland forest soils analysed through canonical analysis of principal coordinates (CAP) computed upon the Bray-Curtis index. Samples have been coloured and shaped by Management. Different sample size was obtained depend on the target taxa → Prokaryotes: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4; Fungi: FOLD=4, FMID=4, FYM=3, CCR=4, CCGA=3, CCS=4 Protists: FOLD=4, FMID=4, FYM=3, CCR=4, CCGA=1, CCS=3; Nematodes: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4; Annelida: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4; Micro-arthropods: FOLD=4, FMID=4, FYM=4, CCR=4, CCGA=4, CCS=4. Significant correlated ASVs (r > 0.3; p.adjusted fdr < 0.1) with LDA scores were represented by arrows.*

The composition of most prokaryotes and all eukaryotic groups studied in the forest soils of West Finland showed a strong response to management, with fungi, nematodes, annelids and micro-arhtropods showing the strongest differences between Continuous cover management (specifically, CCGA and CCS) and Clear cut (specifically, FOLD and FYM). In general, CCGA samples disposition were linked to Agaricales, Rhabitidia and Chromadorina orders; CCS to the Nematodes ASVs from Dorylamida or unlcassifed. For the Clear cut treatments, fungi ASVs from Helotiales and Sebacinales orders and linked to FYM. Moreover, micro-arthropods ASVs from Mesostigmata and Sarcoptiformes order link also to FYM or FOLD samples. The strong impacts of management and the high number of significantly correlated ASVs across most taxonomic groups analyses highlight the impact of management on specific eukaryotic organisms that is difficult to observe in other analyzed agroecosystems. These results will be further analyzed in the upcoming months of the project, combining them with metagenomic data to investigate potential metabolic pathways that are enhanced under Continuous Cover Managements compared to the clear cut ones.

3.4. Analysis of soil biome complexity with co-occurrence networks

We performed co-occurrence network analysis on species abundance data to study the potential interactions within the soil biome. Co-occurrence networks are used to describe how soil organisms potentially co-exist and interact with each other in soil, including both trophic and non-trophic (e.g. antagonism, mutualism, facilitation) interactions. The complexity of networks is described not only by the number of potential interactions among organisms, but also by the type and strength of potential interactions and by the patterns that they form (Guseva et al. 2022). It is thought that the more complex interactions within the soil biomes are, the more likely they are to support soil functioning, even in presence of stressors (de Vries and Wallenstein, 2017). Therefore, understanding how soil biome complexity responds to soil management complements the analysis of soil organisms´ abundance and diversity presented above.

Since we found striking differences in terms of taxonomical diversity across NUTS-2 regions and the way taxonomical diversity responds to soil management, we analyzed changes in soil co-occurrence networks separately for each NUTS-2 region. For each region we generated and analyzed co-occurrence networks observed in soils with contrasting types of management. In this analysis we did not include soil degradation levels, as this would lead to using five replicates for each combination of management type and degradation level within each region, which is not sufficient for constructing reliable co-occurrence networks. We compared alternative and conventional management for West Flanders (Figures 13A, n=10 and 13B, n=10), Murcia (Figures 13C, n=10 and 13D, n=10), Latvia (Figures 13E n=10 and 13F, n=10), South Transdanubia (Figures 14A, n=10 and 14B, n=10), Buenos Aires (Figures 15A, n=10 and 15B, n=10), West Cameroon (Figures 15C, n=10 and 15D, n=10) and Chiangrai (Figures 15D, n=14 and 15E, n=14). For Middle Jutland/South Denmark, we generated three networks: one for fields that have been managed alternatively since at least 3 decades (Figures 14C, n = 10), a second network for fields that have been recently converted to alternative management (Figures 14D, n = 10) and a third network for conventionally managed fields (Figures 14E, n = 10). For Southern Ireland we compared two pairs of networks, namely using data obtained from grasslands with mixture vegetation versus monoculture vegetation (Figures 16A, n=10 and 16B, n=10), and for samples obtained from grasslands with and without agroforestry vegetation (Figures 16C, n=10 and 16D, n=10). Two networks were calculated for continuous cover forests and clear-cutting forests of West Finland (Figures 16E, n=12 and 16F, n=12).

Networks were generated using sequencing data for Prokaryotes (16S region), Fungi (ITS region), Protists (18S V4 region), and Annelids (18S V6V8 region), while we used morphological characterization data for Nematodes and Microarthropods. We used sequencing data resulting from the filtering and 100-fold rarefication, which was applied to all sequencing data for overcoming differences in sequencing depth across samples. We aggregated the ASV-level data at genus level, in order to generate co-occurrence matrices. Starting from genus-level relative abundance data, we calculated the Spearman rank correlation (*rho*) and its the *p*-value for each pair of genera detected in each NUTS-2 region. For this we used the *rcorr* function of the *Hmisc* R package, then p-values were adjusted for false discovery rate (fdr). After that, strong and significant correlations were selected ($rho < -0.9$ and $rho > 0.9$, $p < 0.05$) and were interpreted as network edges, representing connections, or interactions, between genera within the soil biome.

For each network we calculated six metrics that describe in different ways the cohesion or complexity of networks. We quantified the following metrics:

- Number of nodes: the number of genera belonging to that network, i.e. the number of genera connected to at least one other genus.
- Number of edges: the number connections observed in that network.
- Connectance: the proportion of observed connections relative to the total of the possible connections between pairs of genera belonging to that network. This is also known as network

density and ranges between 0 and 1, which are indicative of a poorly connected network and highly connected network, respectively.

- Modularity: it quantifies how strongly genera tend to cluster into modules. Modules are groups of genera that tend to densely connect with each other, but poorly connect with genera belonging to other modules. Modularity ranges between -1 and 1. In our study, high modularity is interpreted as a soil biome characterized by internally cohesive, yet distinct groups of genera. On the contrary, low modularity describes a soil biome that is cohesive in its entirety, which is interpreted as a stabilizing configuration.

- Transitivity: it is a measure of the tendency of genera to form close triangles, or tightly knit clusters. Transitivity, or clustering coefficient is the average probability of two genera connected to a common genus to also connect with each other. This is calculated as the proportion of observed closed triangles relative to all the possible closed triangles in that network. Transitivity ranges from 0 to 1, indicative of poorly and highly clustered networks, respectively.

- Assortativity: it is the tendency, on average, of genera belonging to the same same community of organisms to correlate with each other, relative to their tendency to connect with genera of other communities. Assortativity ranges from -1 to 1, with positive values describing biomes where genera tend to connect within their community, and negative values indicating disassortative mixing, i.e. a tendency of genera to connect with members of other communities.

Since our sample size was relatively low to robustly establish co-occurrence networks, we performed null-models on each one to assess reliability of our estimations. Thus, in addition to calculating the value of each metric for each network, we also estimated the variation of each metric, by randomly permutating the data of each network 1000 times. For each permutation we maintained constant two key characteristics of that network, namely the number of genera and the average number of connections that each genera tends to have with other genera. In this way we maintained the overall number of connections of each random network and evaluate whether the observed patterns of modularity, transitivity and assortativity of the original observed network can be attributed to biologically meaningful patterns rather than to chance, with a 95% confidence level. Network metrics and their estimated standard errors are summarized in Figure 17 and Table S3 in Annex IV.

Figure 13. *Co-occurrence networks in alternatively and conventionally managed soils of West Flanders (A, n=10 and B, n=10), Murcia (C, n=10 and D, n=10) and Latvia (E, n=10 and F, n=10). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists, Nematodes, Microarthropods and Annelids) and by morphological characterization (Nematodes, Microarthropods) from each site.*

Figure 14. *Co-occurrence networks in alternatively and conventionally managed soils of South Transdanubia (A, n=10 andB, n=10), and fields of Middle Jutland/South Denmark managed alternatively (C), recently converted to alternative management (D, n=10) and conventionally managed (E, n=10). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists, Nematodes, Microarthropods and Annelids) and by morphological characterization (Nematodes, Microarthropods) for each site.*

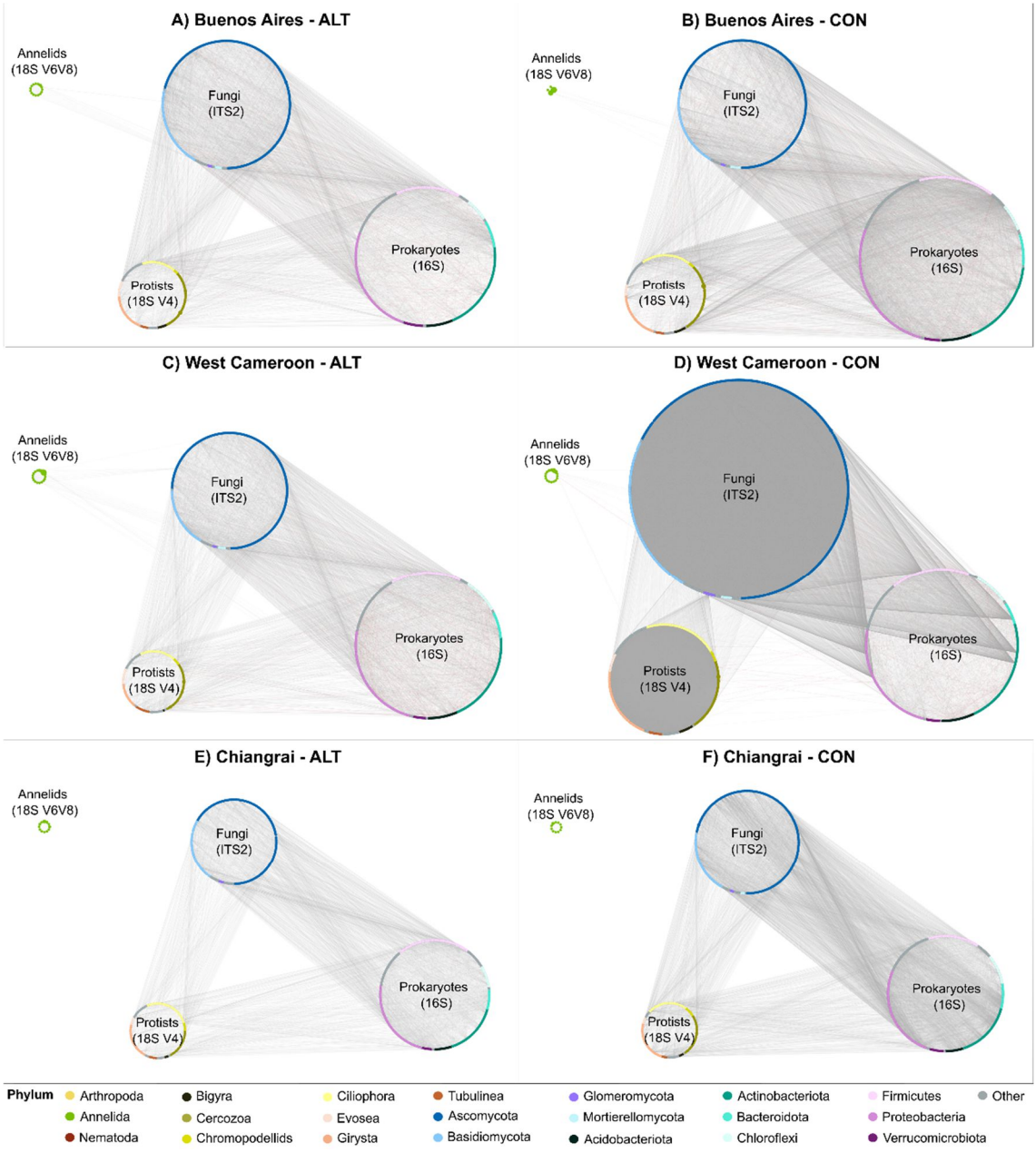


Figure 15. *Co-occurrence networks in alternatively and conventionally managed soils of Buenos Aires (A, n=10 and B, n=10), West Cameroon (C, n=10 and D, n=10) and Chiangrai (E, n=14 and F, n=14). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists and Annelids) for each site.*

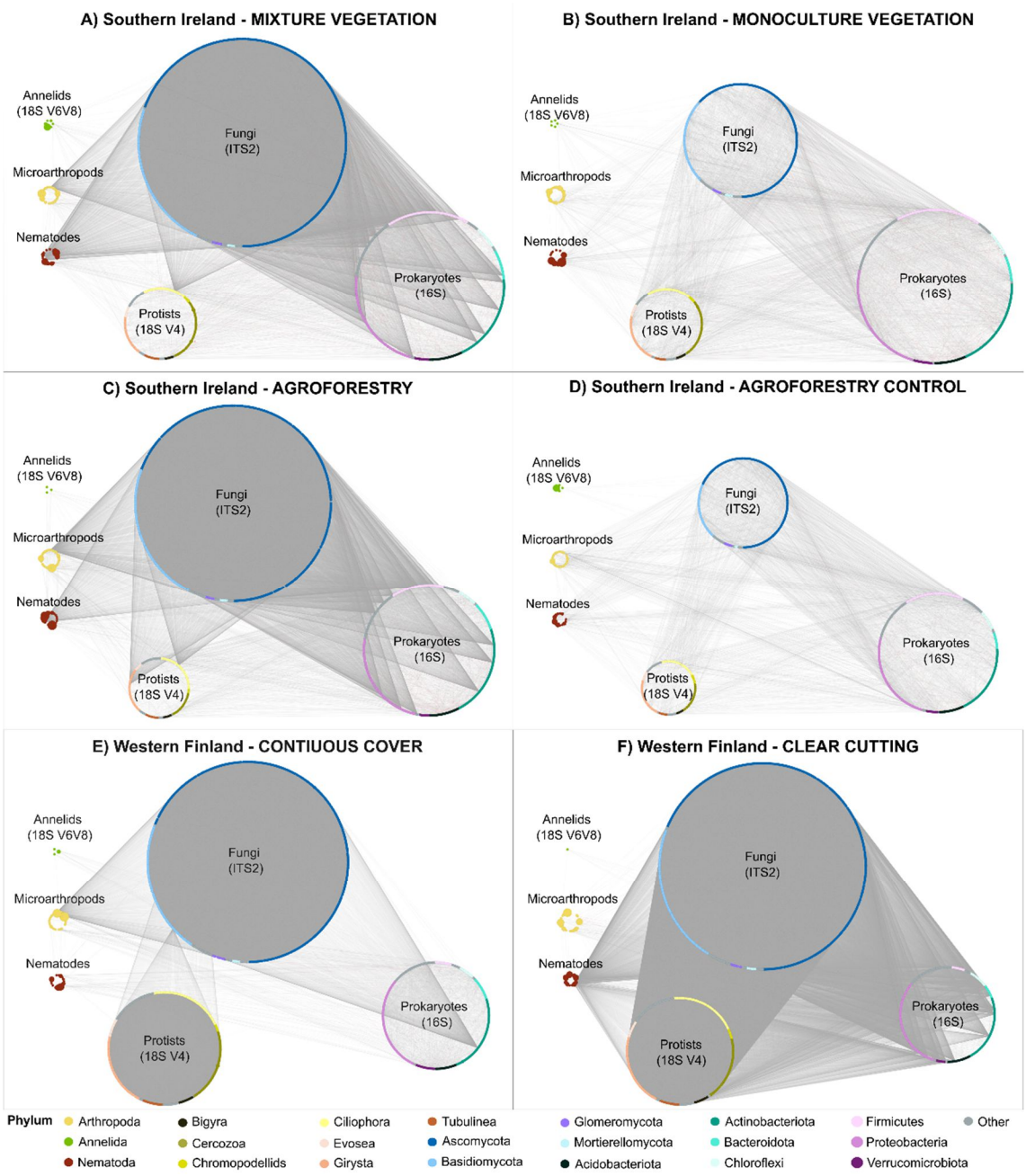


Figure 16. *Co-occurrence networks in alternatively and conventionally managed soils of Ireland pastures with mixture and monoculture vegetation (A, n=10 and B, n=10 and), Ireland grasslands with agroforestry and the relative controls (C, n=10 and D, n=10), as well as forests managed as continuous cover vegetation and with clear-cutting (E, n=12 and F, n=12). Strong positive correlations (rho > 0.9) between pairs of genera are displayed as grey lines, while strong negative correlations (rho < 0.9) are displayed as red lines. All genera connected to at least one other genus are arranged by phylum (marked by colours) within a circle formed by genera belonging to the same community of organisms. The data used for building the networks was generated by amplicon sequencing (Prokaryotes, Fungi, Protists, Nematodes, Microarthropods and Annelids) and by morphological characterization (Nematodes, Microarthropods) from for each site.*

We found that the effect of management types on soil network complexity varied across regions. In soils of **South Transdanubia** we observed a clear effect of management (Figure 14A). There, **alternatively managed soils were characterized by a highly connected network** (16.2% of all possible connections within the network, while conventionally managed soils displayed 1% of them; Table S3 in Annex IV, Figure 17). Alternatively, managed soils of South Transdanubia also had low modularity and high

transitivity, compared to conventionally managed soils. This indicates that the observed genera connected with each other by forming numerous tightly knit clusters, and that such complex interactions were distributed across the whole soil biome, rather than being confined to separate modules.

The assortativity of both networks of South Transdanubia was close to zero, suggesting that organisms tended to form interactions both within and across communities. In this region, we found a strong positive effect of alternative soil management on networks cohesion and complexity of interactions across the soil biome.

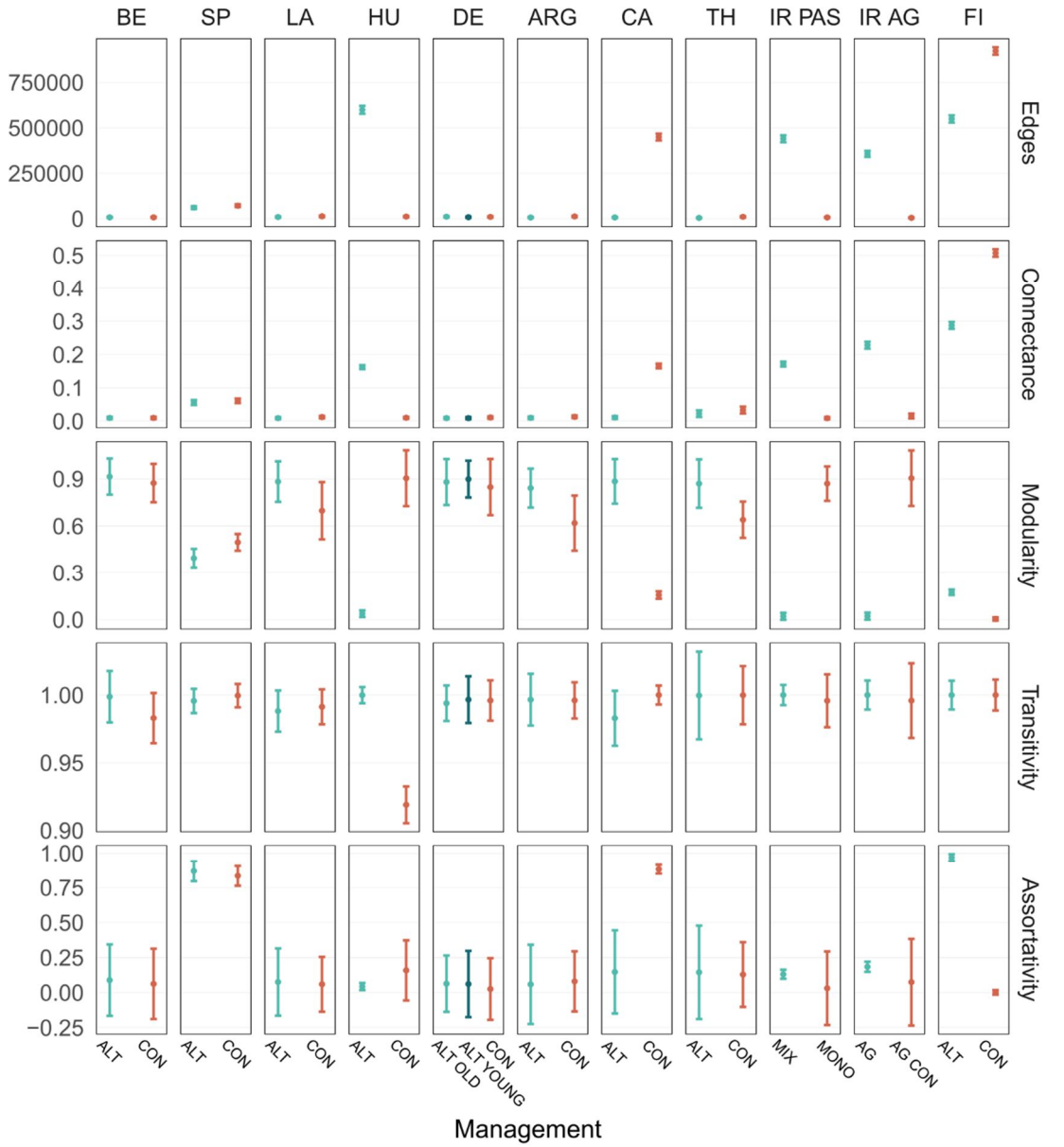


Figure 17. *Network metrics values and standard error represented for each NUTS-2 region - West Flanders (BE), Murcia (SP), Latvia (La), South Transdanubia (HU), Middle Jutland/South Denmark (DE), Buenos Aires (ARG), West Cameroon (CA), Chiangrai (TH), Southern Ireland pastures (IR PAS) and agroforestry sites (IR AG) and West Finland (F) - and management type. The metrics represented are: the number of edges of a network, connectance, modularity, transitivity an assortativity. Standardd errors were calculated with 95% confidence using a null model distribution for that metric, based on 1000 permutations of the observed network. Management is indicated on the x-axis as ALT (alternative management), CON (conventional management), ALT OLD and ALT YOUNG (old and recent alternatively managed soils), MIX (mixture grassland vegetation), MONO (grassland monoculture), AG and AG CONTROL (agroforestry grassland and its control).*

In **Southern Ireland we found a clear positive effect of alternative management on network complexity**, which was found for both the use of mixed vegetation (polycultures) and agroforestry, compared to conventional monoculture grasslands (Figure 16A-D). The use mixture vegetation and agroforestry promoted networks with a high number and density of connections, ranging between 17% and 23% of all possible connections, which was higher than the values of 0.9-1.6% observed in conventional grasslands (Table S3 in Annex IV, Figure 17). Soil networks in alternative grassland vegetation were overall cohesive compared to conventionally managed grasslands, as indicated by the low modularity of networks of mixture pasture and agroforestry compared to those of their respective control soils. This result is particularly relevant since alternative grassland management did not show clear effects on the abundance of soil organisms, which were rather abundant in both alternative and conventional grasslands, however this analysis shows that, within this land-use type, alternative soil management positively affects the complexity of the interactions among soil organisms.

In West Cameroon and West Finland, we found opposing effects of management on network complexity than those observed for Hungary or Ireland. In West Cameroon the network of conventionally managed soils had a high number and density of connections compared to alternatively managed soils, The connections seen for networks of conventionally managed soils were more spread across the soil biome, as indicated by the low tendency of genera to form separate modules, compared to organic management. At the same time the high level of assortativity, close to the maximum value for this metric, indicates that a high proportion of interactions occurred within each community rather than across communities. Figure 15D shows that such a high proportion of homologous interactions were found especially among fungi and among protists. Alternatively managed soils of West Cameroon had overall less complex soil biomes compared to conventionally managed soils.

Networks of forests in West Finland had generally a high number of connections than agricultural soils, and clear-cut forest networks had a higher number and density of connections compared to those of forests with continuous cover (Figures 16E and F). Clear-cut forest networks also had a lower tendency than continuous cover forest networks to form modules and an assortativity close to zero, which characterize their soil biome as overall well connected, both within and across communities.

Finally, in the soils of West Flanders, Murcia, Latvia, Middle Jutland/South Denmark, Buenos Aires, and Chiangrai, we found very similar networks when comparing alternative and conventional management types, indicating that in these biogeographical contexts alternative soil management had limited effects on interactions across the soil biome.

To sum up, we found that alternative soil management had either neutral, positive or negative effects on soil biome complexity, depending on the biogeographical context. In most **arable soils** we found a neutral effect of management on co-occurrence networks, suggesting that practices currently used in organic agriculture might have a limited ability to recover soil biome complexity. In South Transdanubia, however, organic agriculture clearly increased the number and complexity of interactions in co-occurrence networks. Features of organic agriculture specific of this region might in part explain this effect, such as the presence of crop rotation and the limited use of even organic fertilizers. Within this year, we will explore finer differences in management practices in order to elucidate the context-dependent effects of organic agriculture on co-occurrence networks. **In grassland soils**, having a more diverse vegetation had a small impact on the abundance and diversity of soil organisms, which were already relatively high in conventional monoculture grasslands, however it increased the potential interactions among soil organisms. In **forests** we observed an opposite pattern, as clear-cut forests had especially more connections among organisms belonging to different communities, compared to continuous cover forests. In order to interpret this result, we need to include weak interactions in our analysis and to analyze potential trophic interactions.

3.5. Evaluating potential soil biodiversity indicators

With the new EU soil law currently in process, and with an urgent need to develop scientifically-sound methodologies and indicators to monitor soil biodiversity, Soilguard´s WP2 is contributing very relevant information in this regard. Firstly, we used a battery of soil monitoring techniques in a subset of our sites to see which method capture best the diversity and composition of soil faunal groups. Concretely we compare DNA extraction using two contrasting amounts of soil sample, and also compared two contrasting methodologies: mesofaunal extraction and taxonomic identification (microscopic method) vs DNA sequencing. Second, we used our comprehensive assessment on soil biodiversity, and our sampling spanning a wide array of biomes and environmental conditions, to evaluate efficient potential indicators of soil biodiversity for future monitoring schemes.

3.5.1. Comparison between taxonomic and DNA-sequencing estimations of soil faunal composition

While the microscopic method is considered to be the 'golden standard' in soil fauna research, this traditional method is based on a more labour-intensive procedure including taxonomic identification. Molecular methods based on DNA sequencing approaches are emerging as a less demanding alternative in terms of taxonomic expertise. In the SOILGUARD project, we employed both methods for comparison. Various primers have been tested to assess soil organisms in these DNA samples, but for nematodes and microarthropods, two have proven to be the most effective: 18SV6V8 for nematodes and 18SV4 for acari and collembola (a general eukaryote primer).

Both the microscopically obtained data and the molecular data can be used to quantify the soil animals per 100 g soil. However, comparing counted numbers from microscopically identified soil animals with DNA molecular reads data is today still complicated (Geisen et al., 2018; Griffiths et al., 2018). Therefore, using presence/abundance data and relative abundances of nematodes and microarthropods obtained with microscopic and molecular methods, we obtained different values of richness and composition of these faunal groups. In addition, based on the presence/absence data, we calculated the taxonomic similarity between the three types of data, i.e. microscope identification, 18Sv4 and 18SV6V8 molecular DNA data. The taxonomic similarity between methodologies was calculated using the Sörensen similarity index (SI = (2 * number of taxa in common) / (number of taxa in sample 1 + number of taxa in sample 2)*100%) (Sørensen, 1948). Additionally, using the nematode relative abundance, we calculated the Maturity Index for samples derived from both the morphological analysis and the molecular methods (specifically using the nematode-specific primer 18SV6V8). The Maturity Index is the weighted mean of the individual c-p values of the nematodes present in the samples, and was calculated as MI = (∑ (pi x cp-value)) / (∑ pi) (Bongers, 1990; Ferris et al., 2001), where pi is the proportion of a specific nematode taxa and cp-value is their correspondent colonizer-persister value. To calculate the Maturity Index, we uploaded the count table with taxonomic information from the morphological analysis and the count table from the data obtained using the 18SV6V8 primer into the Nematode Indicator Joint Analysis (NINJA) program (Sieriebriennikov et al., 2014, http://sieri ebrie nnikov.shiny apps.io/ninja/ consulted on in May 2024). NINJA was also used to assign nematodes to the colonizer-persister (c-p) scale (from 1 to 5; Bongers, 1990; Ferris et al., 2001).

In total, 69 unique nematode families were detected and 141 genera. Twenty families were detected (by microscope or DNA) in ≥25% of the samples, i.e., these are the most frequently detected taxa (Table 1). The 18S-V6V8 primer detected the highest number of families (n=57) and the 18S-V4 the lowest number (n=41). The microscope detected 47 nematode families. Of these frequently found families, 5 showed a co-occurrence of >50% when comparing the 18S-V6V8 data with the microscope data, and only 2 taxa had a co-occurrence of >80% (first 5 and 2 families in Table 1, respectively; also, when present, these families had high dominance (high %Top5 score)). The 2 most co-occurring taxa were also the 2 most frequently found taxa (present in >85% of the samples). Six taxa had a co-occurrence <20% (the last 6 taxa in Table 1). From the 141 unique nematode genera, 104 were detected using the 18S-V6V8 primer and 87 using the microscope. Only 50 genera were detected by both methods.

Table 1. *Nematodes and microarthropods families occurring in more than 25% of the samples with at least one of the three detection methods: using microscopic identification, molecular detection using the 18S-V4 primer, and molecular detection using the 18S-V6V8 primer. The nematodes are ranked based on decreasing taxa co-occurrence in*

molecularly detected samples (with the 18S-V6V8 primer) and microscopically identified samples. The microarthropod families are given in alphabetic order. %Top5 is the number of times a family detected with microscopic identification was among the five most common taxa relative to the total number of samples in which the family was found.

Nematode Family		**18S-V4**	**18S-V6V8**	**Microscope**	**%Top5**		**Micro-arthropod Family**		**18-SV4**	**18S-V6V8**	**Microscope**	**%Top5**
Cephalobidae	%	81	98	96	89		Isotomidae	%	23.0	42.3	71.3	74
Tylenchidae	%	51	85	88	75		Neanuridae	%	9.2	12.4	28.0	22
Rhabditidae	%	12	66	69	63		Onychiuridae	%	3.4	8.0	28.0	54
Teratocephalidae	%	72	24	17	57		Tullbergiidae	%	2.3	14.6	56.7	52
Aphelenchoididae	%	3	84	63	46							
Plectidae	%	7	50	46	45		# Collembola Families	n	8	11	14	
Pratylenchidae	%	26	56	46	36		# Collembola Genera	n	10	9	46	
Telotylenchidae	%	21	40	54	68							
Aphelenchidae	%	15	40	40	52		Acaridae	%	12.6	26.3	50.0	48
Hoplolaimidae	%	2	28	24	38		Alicorhagiidae	%	2.3	3.6	33.5	22
Prismatolaimidae	%	20	35	15	8		Ascidae	%	0.0	0.0	38.4	63
Heteroderidae	%	23	33	12	26		Digamasellidae	%	0.0	0.0	25.0	37
Tylenchulidae	%	21	37	34	52		Eupodidae	%	5.7	13.9	63.4	52
Anguinidae	%	25	25	58	40		Nanorchestidae	%	0.0	0.0	42.7	50
Alaimidae	%	0	32	20	30		Oppiidae	%	20.7	1.5	43.9	60
Diplogastridae	%	3	23	45	47		Parasitidae	%	2.3	16.1	32.3	17
Merliniidae	%	11	60	10	63		Pygmephoridae	%	50.6	0.0	0.0	-
Qudsianematidae	%	0	15	75	47		Rhagidiidae	%	12.6	45.3	32.9	13
Monhysteridae	%	1	4	33	31		Rhodacaridae	%	0.0	0.0	50.6	76
Nordiidae	%	4	40	0	-		Scutacaridae	%	0.0	0.0	46.3	49
							Tectocepheidae	%	14.9	18.2	28.0	30
# Nematode Families	n	41	57	47			# Acari Families	n	28	34	74	
# Nematode Samples	n	151	163	164			#Acari Genera	n	24	28	48	

* Occurring in ≥25% of the samples for at least one of the detection methods. ** Nematodes were analysed separatel, Collembola and Acari were analysed together. *** Calculated for only the microscope samples.

In total, 16 Collembola families were detected, 14 using the microscope, 11 using the 18S-V6V8 primer and only 8 using the 18S-V4 primer (Table 1). The 18SV6V8 primer detected 2 families that were not detected by the other methods. Four families were detected in ≥25% of the samples (all 4 with the microscope, only 1 with 18SV6V8, and 0 with 18SV4). In total, 50 Collembola genera were detected of which 46 with the microscope and only 9-10 using DNA.

In total, 80 Acari families were detected, of which 74 with the microscope, 34 with 18S-V6V8 and only 28 with the 18S-V4 primer (Table 1). Of these, 13 families were detected in ≥25% of the samples, 12 with the microscope and only 3 using DNA. Of the 13 frequently detected families, 6 were not detected with the 18S-V6V8 primer and 5 not with the 18S-V4 primer. Primer 18S-V4 detected one frequently found family that was not detected by the other two methods. Only 76 genera were detected of which 48 were found with the microscope and only 28 with the 18S-V6V8 and 24 with 18S-V4 primer. The primers detected 28 genera that were not detected with the microscope.

To further compare the (dis)similarity in the detection of nematode and microarthropod taxa (in this case Collembola and Acari together) using microscope or DNA-sequencing estimations, figures 18 and 19 present the taxonomic community similarity based on presence-absence of families for all sampling sites. When comparing the two primers for the nematode data, only 5% of the samples had a similarity >80% (Figure 18). 38% of the samples had a similarity <50%. When comparing the primers with the nematode microscope data, these scores were lower. The highest similarity between the 18S-V6V8 primer and the microscope data was 74%, and 80% for the 18S-V4 primer. As much as 45% of the 18S-V6V8 samples had a similarity <50% with the microscope samples, and this was even as high as 70% in the case of the 18S-V4 primer.

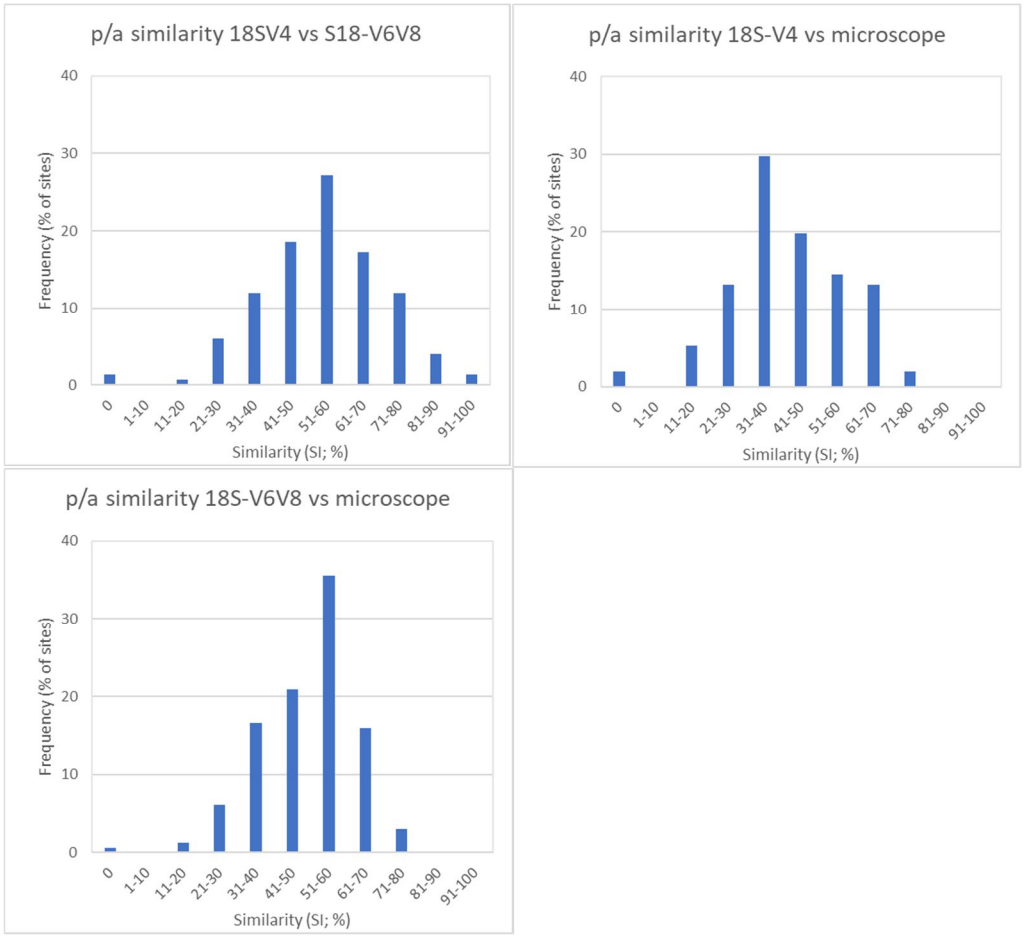


Figure 18. *The frequency distribution of the Sorensen similarity index of nematode family taxonomic composition for the European SOILGUARD samples analysed microscopically and molecularly (eDNA) using two different primers (18SV4, 18SV6V8). The European SOILGUARD samples included Spain, Hungary, Denmark, Hungary, Ireland, Latvia, Belgium and Finland, and 151 samples were available after rarefaction of the molecular data obtained with the 18SV4 primer, while 163 samples were available after rarefaction of the molecular data obtained with the 18SV6V8 primer. Left above panel: Similarity index for the presence-absence data for 18SV4 versus 18SV6V8 primers data (n=151), right above panel: Similarity index for the 18SV4 primer data versus microscope data (n=151), and left below: Similarity index for the 18SV6V8 primer data and microscope data (n=163).*

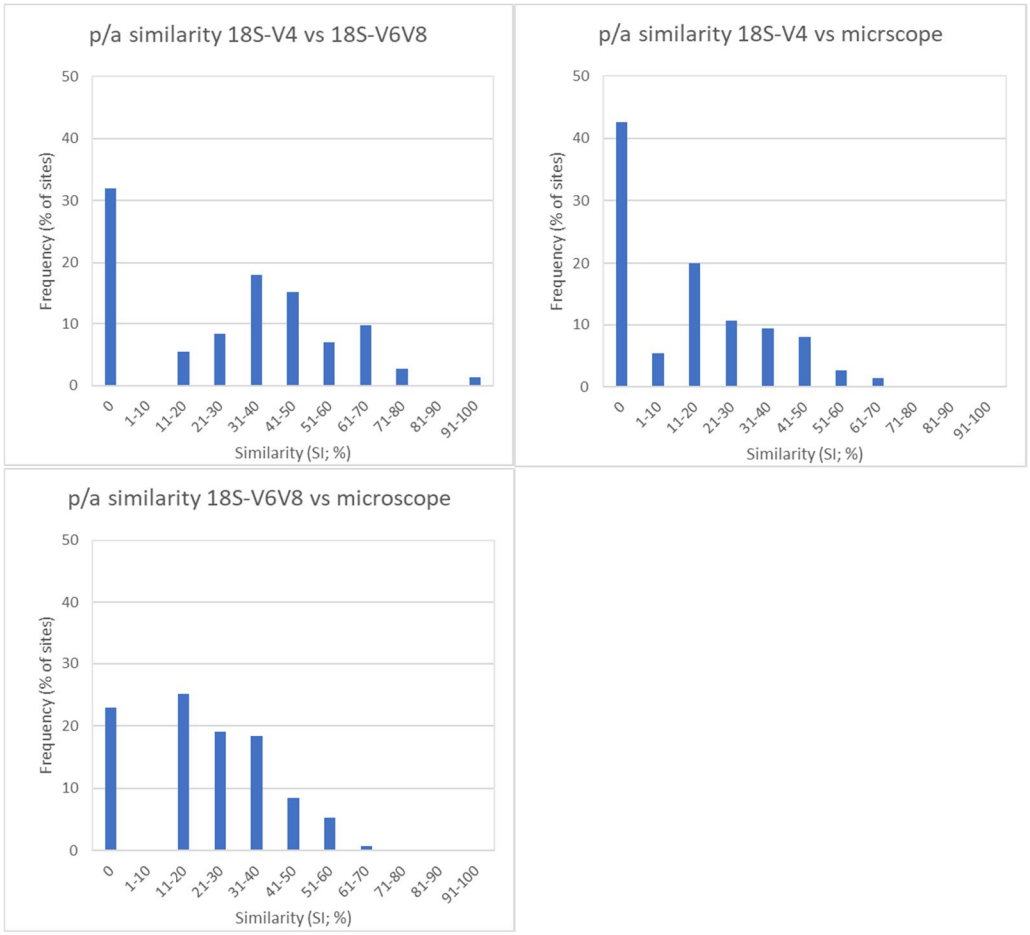


Figure 19. *The frequency distribution of the Sorensen similarity index of micro-arthropod family taxonomic composition for the European SOILGUARD samples analysed microscopically and molecularly (eDNA) using two different primers (18SV4, 18SV6V8). The European SOILGUARD samples included Spain, Hungary, Denmark, Hungary, Ireland, Latvia, Belgium and Finland, and 75 samples were available after rarefaction of the molecular data obtained with the 18SV4 primer, while 131 samples were available after rarefaction of the molecular data obtained with the 18SV6V8 primer. Left above panel: Similarity index for the presence-absence data for 18SV4 versus 18SV6V8 primers data (n=75), right above panel: Similarity index for the 18SV4 primer data versus microscope data (n=75), and left below panel: Similarity index for the 18SV6V8 primer data and microscope data (n=131).*

In the case of the microarthropod samples, a high proportion of the samples did not share any family (0% similarity) when using different methodologies (Figure 19). When comparing the two primers, 4% of the samples had as similarity >80%. As many as 79% of the samples had a similarity <50%. Comparing the primer data with the microscope data showed even lower levels of taxonomic similarity. 94% of the 18S-V6V8 samples had a similarity of <50% with the microscope data, and this was even 96% for the 18S-V4 primer. For both the nematodes and micro-arthropods, the presented families in Table 1 focused on the most frequently detected families. Most of these families also appeared to be among the top-5 most abundant taxa sites where they were detected (Table 1).

We found a positive relationship between the number of Acari and Collembola families detected in a location (comparison based on microscopic data) (Figure 20). Finnish forest soils were the richest in Acari families (>15 families per site) but did not differ in Collembola family richness.

In addition, we found that the relationship between the nematode-based indicator Maturity Index calculated with data from the morphological analysis, and data from the molecular analysis done with the best yielding primer (18SV6V8) had a very low r squared (R^2=0.054) (Figure 21). This indicates that the relationship between the same indicators calculated with data from different analysis is very low, and one type of data cannot explain the others and therefore used interchangeably.

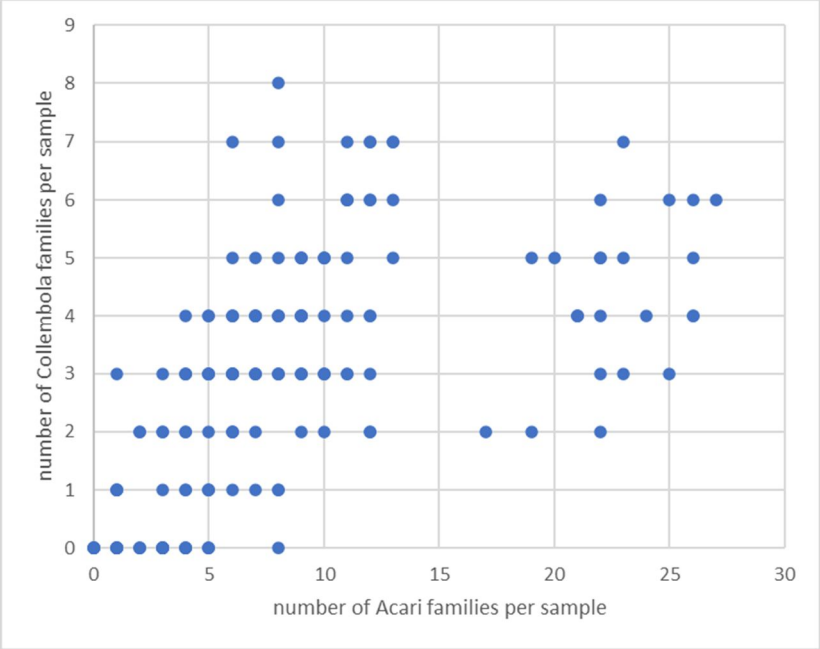


Figure 20. *Correlation between the Acari and Collembola family richness both calculated based on microscopic observations in the 164 SOILGUARD European sites (Spain, Hungary, Denmark, Hungary, Ireland, Latvia, Belgium and Finland). The sites with >15 Acari families were all sampled in Finnish forest soils.*

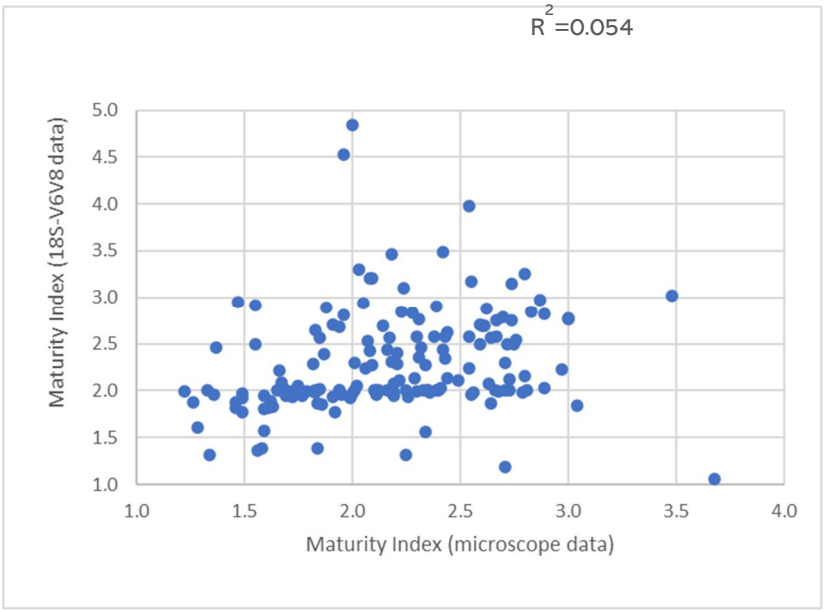


Figure 21. *Relationship between the nematode-based indicator Maturity Index obtained by the NINJA program by using the count table with taxonomic information based on the morphological analysis, and the count table table with taxonomic information based on the data obtained with the primer 18SV6V8. The r squared indicates the coefficient of determination, determining the proportion of variariance that is explained by the explanatory variable.*

Our analyses show high levels of disagreement in DNA data between the two selected primers (18S-V4 and 18S-V6V8). Also, the similarity in taxon detection and taxonomic community composition

between the 18S primers and the microscopic data was very low. One of the reasons for this discrepancy could be the soil volume used to extract nematodes or nematode DNA. For microscopic detection 100 g soil (obtained from a larger composite sample) was used to extract nematodes, while in case of the molecular analyses, sub-samples of only 0.25 g soil were used. The average number of nematodes per 100 g soil was 2242 (excluding the forest samples from Finland). This would imply that 0.25 g soil would represent 5-6 nematode specimens. Furthermore, there are indications that currently the best results of nematode community analyses using molecular detection methods are obtained when the nematodes are first extracted from soil (e.g. from 100 g soil) and then used for DNA analyses (Wiesel et al., 2015).

The current analyses focused on presence-absence data only, i.e. they focused on the detection of specific taxa in the various soil samples. For the calculation of nematode- or microarthropod-based functional group indices, such as feeding groups or life history/tactics groups, taxonomic group numbers per unit soil or relative abundances are required. Counts obtained by microscopic observations can easily be converted into such absolute or relative abundances. Such values are widely used in soil ecological research focusing on, for example, soil food webs (Du Preez et al., 2022). The similar quantification of relative abundances of soil animals based on molecular data is currently still problematic (relative abundances obtained with molecular methods might not coincide with real relative abundances) (Geisen et al., 2018; Griffiths et al., 2018). We can, in fact, see that in the case of nematodes, the Maturity Index calculated with the data obtained with the molecular analysis (despite using the best yielding primer) did not have a relationship with the Maturity Index calculated with traditional 'golden standard' microscopy method.

3.5.2. Identifying key diversity indicators to assess soil biodiversity in EU croplands

Identifying and studying effective soil biondicators is important for understanding and protecting soil biodiversity (Pulleman et al. 2012). As part of the SOILGUARD project, soil biodiversity has been analyzed using a holistic approach across different regions and agroecosystems. Although this wide range of generated indices is essential to fully understand the complexity of biodiversity dynamics for each group of organisms and their possible links with soil functioning, it is impractical for farmers to apply all these methodologies in the field to assess their biodiversity status. For this reason, the work presented in this section focuses on generating clusters of highly correlated biodiversity indices. The main aim of this is to suggest the most straightforward methodologies to monitoring biodiversity that help to answer the specific question of farmers or directives.

Beforehand, we need to emphasize that this answer is far from simple, considering the complexity of soil communities and the relatively small sample size of our study. We consider this work as a first step to define soil bioindicators for biodiversity within the project, but there is a need for wider datasets and experimental designs to increase the robustness of results. However, we will try to counteract the small sample size by first analyzing the links between the correlations of croplands in Europe (presented in this report), using the international regions and the non-crop biomes as an out-of-sample test (not shown in this report, work under progress).

For the first step, we selected all EU cropland sites including both alternative and conventional management. These encompass a total of 100 sites: 20 West Flanders (BE), 20 Murcia (ES), 20 Middle Jutland / South Denmark (DE), 20 Latvia (LV) and 20 South Transdanubia (HU). Samples from young alternative management of Denmark (n=10) were removed to have a balanced dataset between regions. Next, we identified connections not only between the diversity indices themselves but also with other measurements that can indicate ecological processes in the soil, including information on enzymatic activities, leaf damage as an indirect measurement of pathogen abundance, genes related to the nitrogen cycle, and the presence of nitrogen in different forms in the soil (Supplementary Table S4, Annex III).

Before starting the analysis, we excluded variables with non-numeric values in more than 10% of samples of the database. Thereafter, Spearman´s correlations between biodiversity and ecological processes were performed using corr.test function from psych R package (V2.2.5), considering correlations

with r > |0.5| as strong related variables. Variables showing the higher number of significative correlations will be considered as potential good indicators for biodiversity in the soil.

After generating the correlation matrix, we identified two clusters of biodiversity measurements and another cluster of nitrogen related variables with high significant correlations amongst them (Figure 22). From left to right, the first cluster is based on data from classic microscopic techniques (acari and collembola richness/abundance) and PLFA quatification per gram of soil (methanotrophs, fungi, actinobacteria, bacteria, and tmb). The second cluster is characterized by the connection between the quantification of nitrogen-cycling genes and the different forms of nitrogen in the soil. The third cluster encompasses all variables related to richness measured by DNA sequencing techniques (16S, ITS, 18S V4, 18S V4V5, and 18S V6V8).

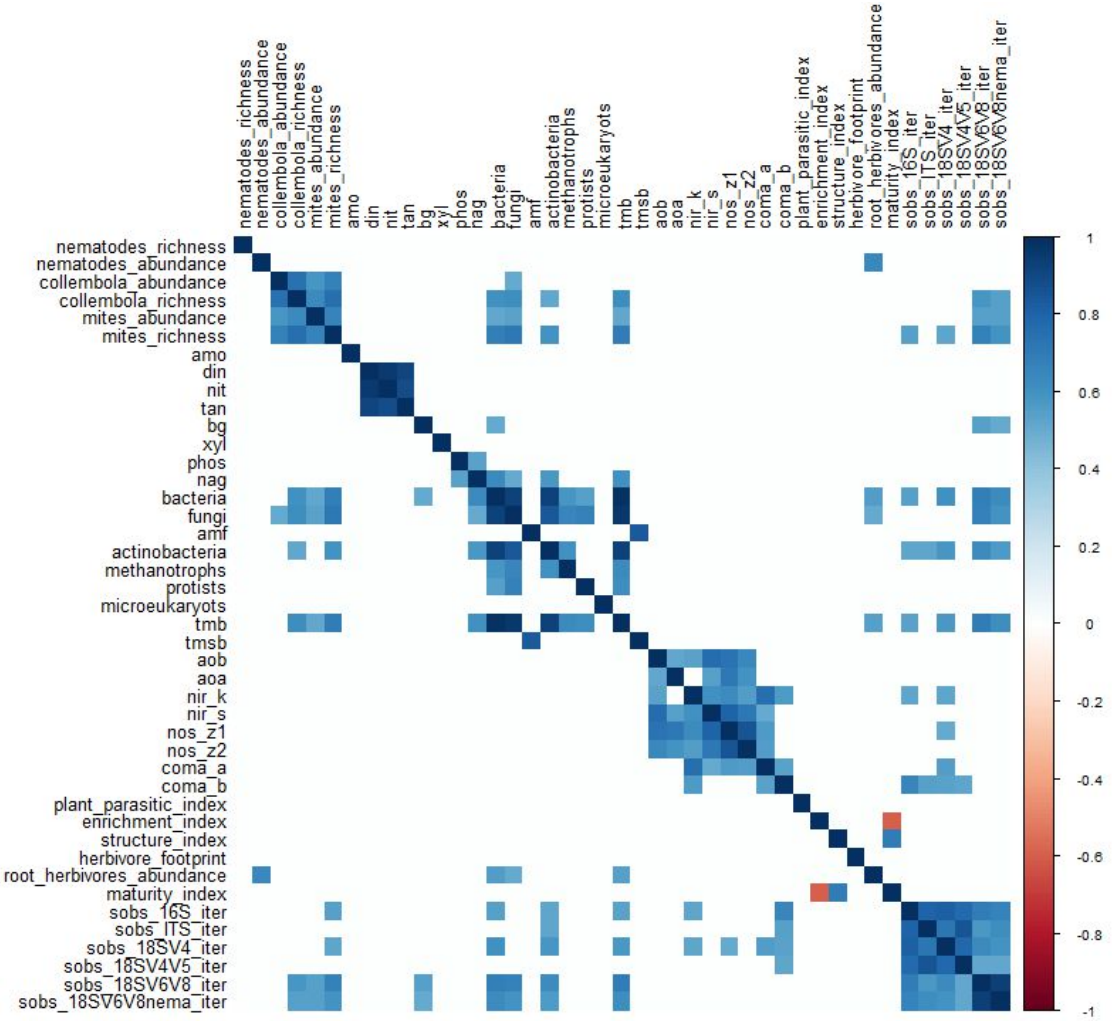


Figure 22. *Biodiversity indicators correlation heatmap. Cells were coloured based on the correlation matrix, with blue gradient when closer to +1 and red when closer to -1, with correlations higher than 0.5, are coloured in blue when closer to +1 and red when closer to -1. Low correlations (r> |0.5|) were removed and represented as blank cells.*

In a second step, we conducted a preliminary analysis including the most prevalent species (assessed by sequencing methodologies) in the same sites, to the previous dataset of biodiversity and ecological processes measurements. For this, we select ASVs present in more than 50% of the samples and with a total relative abundance greater than 5% across each sequencing dataset. Thus, we retained a total of 398 ASVs from 16S, 18S V4, 18S ITS, and 18S V6V8 (the generation of these datasets is explained in Annex III). The 18SV4V5 and 16Smit datasets were not used in this case because they had non-numeric values in more than 10% of the cropland related samples. After performing multiple Spearman

correlations, the top 10 ASVs or indices with higher number of significative interactions were identified and discussed.

The top 10 variables with more significative interactions among all biodiversity features are represented in Table 2. In terms of richness derived from sequencing, ASVs are observed from global eukaryotes (sobs_18SV4_iter), fungi (sobs_ITS_iter), prokaryotes (sobs_16S_iter) and nematodes (sobs_18SV6V8nema_iter). Among laboratory measurements, bacteria, indirectly measured by PLFAs (bacteria), is the most strongly correlated variable.In terms of species, 2 ASVs belonged to the KD4-96 class (genus: unclassified), 1 from Micrococcales (genus: *Lapillicoccus*) and 1 from Microtrichales (genus: *Illumatobacter*) in the prokaryotes dataset, while 1 belonged to Mortierellales (genus: unclassified) in the fungi dataset.

Table 2. *Top 10 biodiversity index and ASVs analysed in WP2 based on the number of significative correlations among them.*

Nº significant interactions	variable	Taxa
88	sobs_18SV4_iter	-
65	ASV26_16S	p_Chloroflexi; c_KD4-96;
62	ASV223_16S	p_Chloroflexi; c_KD4-96;
62	ASV35_ITS	p_Mortierellomycota; c_Mortierellomycetes; o_Mortierellales; f_Mortierellaceae, g_unlcassified
62	sobs_ITS_iter	-
61	ASV168_16S	p_Actinobacteriota; c_Actinobacteria; o_Micrococcales; f_Intrasporangiaceae; g_Lapillicoccus
58	bacteria	-
58	sobs_16S_iter	-
58	sobs_18SV6V8nema_iter	-
57	ASV86_16S	p_Actinobacteriota; c_Acidimicrobiia; o_Microtrichales; f_Illumatobacteraceae; g_Illumatobacter

To identify the main explanatory axis of the biodiversity variation, we merged the previously mentioned 398 ASVs, which can be considered as “common species”, and calculated a new dissimilarity matrix based on Bray-Curtis distances. After calculating the main explanaroty axis through a PCoA, we assessed if any of the previous biodiversity indices and ecological processess measured (keeping only one if several are highly correlated), including also the soil physiochemical measurements / climate data collected, fit better into these axes, using envfit R function. Correlations were considered significant when $fdr < 0.1$ and $r > |0.5|$ for biodiversity, environmental and ASV parameters. The main aim of this third step is to: i) evaluate if the previous bioindicators selected reflect well environmental changes occurring to soils, and ii) identify ASVs that best indicate community composition changes and links with biodiversity indices.

Although the two main axes of this dataset represent almost 30% of the variability, 14 ASVs and 35 biodiversity index/environmental variables are correlated to the Eigen scores (Figure 23). Going from left to right in Figure 23, we can differentiate the similar distribution of the ASV36_ITS and ASV45_16S, with the richness indices of fungi, collembola, and nematodes, or the inverse of soil infiltration rate. The richness of mites, prokaryotes, and ammonia-oxidizing Nitrospira clade B quantification (coma_b) are

related to sand content. Moreover, the distribution of ASV26_16S, ASV11_16S, and ASV20_16S is related to the richness of global eukaryotes, principal component of prokaryotes beta-diversity, and several measurements of precipitation (pwaq, pweq, and pwm) or ecosystem production.

In contrast, the ASV59286_16S and ASV168_16S fit into the principal component of fungi composition, annual precipitation, and phosphatase activity. The ASV1_ITS fits with precipitation of the driest month and quarter (pdm, pdq), and the ASV520_ITS with precipitation of the coldest quarter (pcq) and available phosphate (available_p). Conversely, the only ASVs from eukaryotes (ASV4_18SV4/24694_18SV4) and the ASV51811_16S have no similaritieswith any biodiversity nor environmental variable but follow the opposite direction compared to the principal component of nematodes beta-diversity and temperature seasonality. Annual temperature range and precipitation seasonality follow a similar distribution with no association to biodiversity features. Temperature seasonality (tse) seems to be related with principal component of nematodes composition. The ASV36_16S fits into the principal component of eukaryotes diversity and ASV49_16S into the mean diurnal range. Mean temperature of the driest quarter and annual (mtdq) fit in the same direction as clay content.

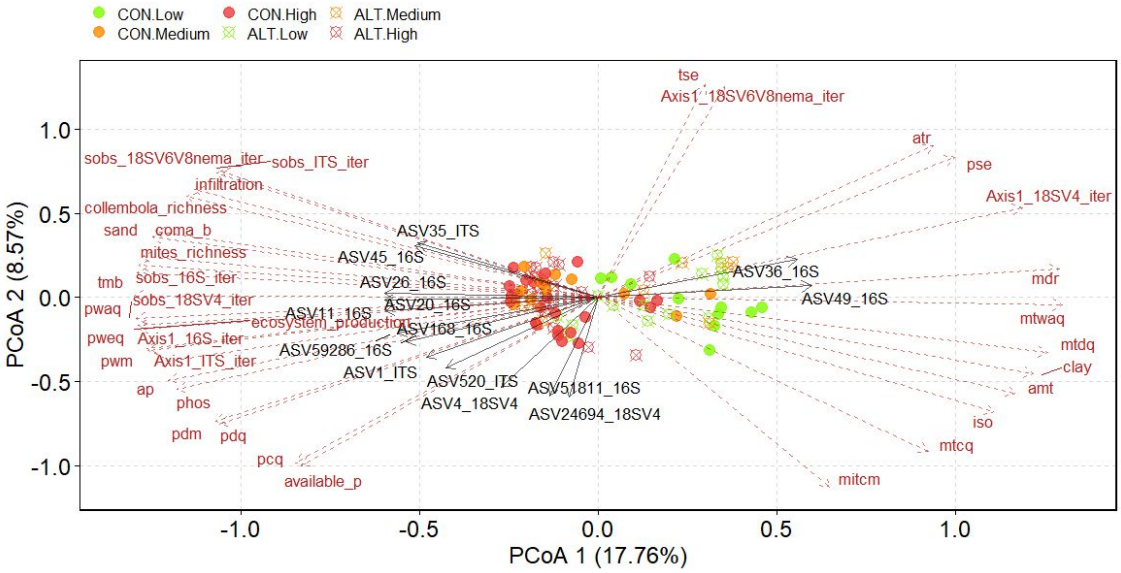


***Figure 23**. Beta-diversity Principal Coordinate Analysis (PCoA) calculated from most prevalent species (common species) of the sequencing datasets, across the cross-biome network sites belonged to croplands from EU regions. Significant correlated ASV and biodiversity or environmental variable (r > |0.5|; p.adjusted fdr < 0.1) with PCO scores were represented by black and red arrows, respectively. Abbreviations Supplementary Table S4 in Annex III.*

To sum up, the current first approach of soil biodiversity indicators suggest, for now, that **the combination of prokaryotes richness observation through 16S rRNA metabarcoding or total microbial biomass, in combination with a quantification of a target gene related with the nitrogen cycle** (aob, oa, coma_a or coma_b) **may be good soil biodiversity bioindicators** because of their good correlation with the rest of the 50 diversity measures in our dataset. When we refer to the ASVs, a significant number of them have been identified, most of them prokaryotes, but also including fungi and other eukaryotes. Although the results reported here are a first sound step to produce reliable and efficient biodiversity indicators, there is still much work to do in this regard. Furthermore, they will serve as a foundation for upcoming analyses of the WP3 data already generated on the impact of climate change on the biodiversity and multifunctionality of agroecosystems. By the end of this year, we expect to have: i) identified the number of “clusters” of correlations between biodiversity indicators, ii) one reliable and cost-effective indicator representing variation on each one of those clusters, and iii) an out-of-sample (using the rest of our WP2 data) evaluation to assess the effectiveness of the developed soil biodiversity indicators to properly assess soil biodiversity status.

4. Soil biodiversity potential to deliver soil multifunctionality

This chapter aims to elucidate the soil biodiversity potential to deliver soil multifunctionality by means of i) defining the role of soil degradation and soil management strategies in the conservation or loss of key soil multifunctionality; and, ii) quantifying the weight of soil biodiversity in the delivery of each of the NCP assessed.

To achieve this, we first conducted a "winner/loser" analysis using our functional indicators, to observe the response of each soil organism under conventional versus alternative soil management (RII). We then employed linear models, similar to those described earlier, to assess soil multifunctionality in relation to a broader set of predictors. In this analysis, we included biodiversity (standardised average of the analysed groups) as one of the predictors. Finally, we aimed to thoroughly evaluate the effectiveness and implications of the EU's Farm to Fork strategy, which seeks to convert 25% of EU agriculture to organic farming by 2030, considering not only soil functionality but also the potential for maximising soil biodiversity, crop yield, and minimising nitrogen loss.

4.1. Measurement of soil properties and functions

We identified the role in the conservation or loss of key soil functions, as defined in collaboration with stakeholders, and linked to important contributions they provide to people (NCP, Nature's Contribution to People) (Table 3). These functions were categorised into various groups according to the NCP provided: food production, soil formation and protection, regulation of hazards and extreme events, regulation of detrimental organisms, and regulation of freshwater quality. The importance of these NCPs was evaluated as a percentage, based on feedback from local stakeholders in the agroecosystems studied (information obtained from questionnaries and analyses developed in WP4).

Table 3. *Soil physico-chemical analyses and other measurements, organized by their relationship with nature contributions to people (NCPs**).** NCPs follow IPBES terminology, and have the valuation (adding up to 100%) by the local stakeholders, according to IÖW, averaged across the different regions. Justification for the relationship with NCPs or other additional information is provided as "comments". LSTs = local sampling teams (partners of each region), rest of acronyms are the institutions abbreviations, as commonly used in Soilguard. Different shading is added for visualisation purposes (to differentiate between different NCPs).*

Related NCP	Measurement	Units	Comments
None, basic physico-chemical attributes	Texture	% sand, clay and silt	General soil characterization, not directly related to NCPs, but important to interpret results
	Bulk density	g soil/cm^3 volume	
	pH	Unitless	
	Electric conductivity	microS/cm	
Food production (**32.1%; range 18.6-44.5%**)	Crop yield	Kg/Ha	Obtained directly from the farmers or measured in situ by LSTs
	Stability in crop production	Unitless	Calculated as the reverse of CV of NDVI mean.
Soil formation and protection (**22.6%; range 13.8-27.4%**)	Soil aggregates stability	Semi-quantitative (scores 1 to 12)	Related to resistance to further erosion
	Available P	mg P/kg soil	Plant-available nutrients
	Available N	mg NO3-/kg soil + mg NH4-/kg soil	
	Litter decomposition (tea bag index)	% weight loss/day	Potential capacity of the soil to degrade organic matter
	Soil enzymatic activities (beta-glucosidase)		

	Soil enzymatic activities (xylanase)	nano-mols of MUF (methylumbelliferyl)/g. dry soil · hour	
	Soil enzymatic activities (N-acetylglucosaminidase)		
	Soil enzymatic activities (acid phosphatase)		Potential capacity of the soil to obtain P
	Potential N mineralization	mg N/kg soil · day	Potential capacity of the soil to transform ammonia to nitrate
	N transformation rate	mg N/kg soil · day	Capacity of the soil to transform different forms of N (N cycling)
	Depolymerization	mg N/kg soil · day	Release of organic N monomers into the soil
	N cycle genes	Abundance of AmoA/AmoB genes	Quantifies different pathways of the N cycle
	Amount of mycorrhizal fungi	Estimated biomass of AMF, based on NLFAs	Mycorrhizal fungi aids crop growth, mainly under nutrient-limiting conditions
Climate regulation (**21.6%; range 14.7-34.5%)**	Soil organic C	g. organic C/kg soil	Will be combined with bulk density to obtain soil C stocks
	Methanotrophs	Estimated biomass of methanotrophs, based on PLFAs	Improve the balance of greenhouse gas exchanges in the soil
Regulation of hazards and extreme events (**6.1%; range 1.1-10.3%)**	Water infiltration	Amount of time in infiltrating 50% of the 10 ml added	Helps with flood and drought regulation
	Water holding capacity	% (g. water retained/g. dry soil)	
Regulation of detrimental organisms (**5.3%; range 1.9-10.6%)**	Leaf damage	% of leaf surface damaged by pathogenic fungi or herbivorous insects	Estimated from leaf pictures
	Root feeding nematodes	% of nematodes that are herbivores, multiplied by the amount of nematodes/100g	Control insects population contributing to the general crop health and productivity
Regulation of freshwater quality (**12.2%; range 8.4-20.7%)**	NO3/PO3 in leachates	mg N/ L of leachate, mg P/ L of leachate	Nutrient retention capacity resulting from fertilization

Soil texture was analyzed by measuring the proportions of sand, silt, and clay following Kettler et al. (2001). For the analyses presented here, sand content was chosen as the variable of interest due to its high correlation with silt and clay. Soil pH and conductivity were measured using a pH meter with a 1:5 soil-water extraction. Soil bulk density was measured in situ for each plot using the cylindrical core method (Blake and Hartge, 1986).

We determined water-holding capacity (WHC), of each soil sample using a subsample of 20 grams of dry soil that was saturated by 20 mL of deionized water, covered and allowed to drain for 24 hours (Juárez et al., 2004). To measure soil infiltration, 20 grams of soil were placed in a funnel and saturated with water for one hour, covered with a plastic film. Then, 10 mL of deionized water was added to each saturated soil sample, and the time taken for 50% of the water to pass through was recorded (adapted from Mills et al. 2009).

Soil organic carbon, was obtained by ^{13}C isotopic analyses after acid fumigation (Harris et al., 2001), which avoid the use of the highly pollutant Potassium dichromate and adheres to EU´s green card standards. Soil nitrogen transformation rates (i.e. net potential transformation of N, mineralization, ammonification, nitrification, and depolymerization) (Schimel and Bennet, 2004) were determined based on the measurement of total extractable nitrogen (TAN) by KCl, available mineral N (NH_4^+-N and NO_3^--N), and dissolved organic nitrogen (DON) before and after soil incubation in the laboratory at 80% field capacity and 30° C for 14 days (Allen, Grimshaw, and Rowland, 1986).

Extractions were performed using 1M KCl at a 1:10 ratio. Conversion of DON to NO_3- was achieved through autoclave digestion (Sollins et al., 1999), NO_3^- to NH_4^+ conversion was done by Devarda alloy reduction, and NH_4^+ determination was conducted using the indophenol blue method (Sims et al., 1995). Available P was measured by extraction with 0.5M $NaHCO_3$ (pH 8.5) (Olsen et al., 1954). Soil extracts were agitated for 16 hours and then centrifuged at 900 g for 30 minutes (Guppy et al., 2000). The concentration of PO^{3-}_4 -P in the supernatant was used to estimate Pi content by colorimetry (absorbance at 660 nm wavelength), following the malachite green method (Fernández et al., 1985; modified from Hess & Derr, 1975).

We determined the soil enzymatic activity to determine the nutrient degradation capacity of the soils. Soil enzyme activities were measured using a microplate fluorometric assay (Dick et al., 2018), based on the detection of methylumbelliferone (MUF) released by enzymatic hydrolysis of specific substrates when incubated with soil at the optimal pH for the enzyme tested. Enzyme activity was determined by measuring the amount of *p*-nitrophenol (PNF) released from 0.5 g of soil after incubation at 37°C for 1 hour with different substrates in MUB buffer (pH 6.5).

We assessed litter decomposition in response to specific nutrient inputs into the soil. Litter decomposition was quantified using the Tea Bag Index (TBI). The bags were buried at a depth of 8 cm within the active soil layer for approximately 45 days. Afterward, they were subjected to a 48-hour drying process in an oven at 70°C, followed by weighing to calculate the litter decomposition rate (k) (Keuskamp et al., 2013 for additional details).

The short-term potential of N and P loss by leaching was performed in a laboratory experiment in which precipitation was manipulated. 30gr of dry soil were weighted and added into two percolation columns per sample. One of them were fertilized with 1ml of a dissolution of KNO_3 and KH_2PO_4 with an application rates of 200 mg N g-1 soil and 1000 mg P g-1, and distilled water was added to reach 60%WHC. The other percolation column was used as control, without the fertilized solution. The incubation was carried out at 25°C in the dark. After 5 days of incubation, a precipitation of about 50L m^{-2} was simulated by adding 50 ml of distilled water to each column with the stopcock closed to improve mixing between soil and water. After 5 minutes, the stopcock was opened, and the leachate was collected. Ammonium, nitrate and phosphate contend were determined in leachate simples after filtration, and results are expressed as mg N-retained g^{-1} soil and mg P-retained g^{-1} soil (Wang et al., 2021).

To determine soil erosion resistance, we assessed the physical and chemical stability of aggregates. The procedure consists of 3 phases. Firstly, the slake phase determines physical resistance to wetting after immersion in distilled water (values from 0 to 4, indicating no fragments to aggregate remaining intact, respectively). Additionally, chemical resistance (dispersion) of fine soil particles is determined (values from 1 to 4, in increasing order of sodification). Lastly, the remould phase relates to those aggregates that showed no dispersion in the previous phases, with values following the same scale as before. The sum of these values provides the erosion resistance degree, ranging from 0 (low resistance) to 12 (fully resistant to soil erosion) (Field et al., 1997; Tongway and Hindley, 2004).

We visually estimated the percentage of leaf damage caused by both herbivorous insects (i.e., chewing damage, sap-sucking or rasping damage, leaf-mining damage and galling) and pathogenic fungi [i.e., rust fungi, downy mildews, powdery mildews and leaf spots]. (See full protocol at https://www.bug-net.org/detailed-protocol-estimating-leaf-damage/).

We determined crop production (Kg/Ha) data through surveys conducted with farmers responsible for the crops. Additionally, we estimated ecosystem stability by calculating the reverse of the coefficient of variation (CV) of the mean Normalized Difference Vegetation Index (NDVI) calculated from satellite imaging for each study site (see Annex V- Introducing the Sentinel-2 dataset & the Remote Sensing Indices).

4.2. Calculating soil multifunctionality

To avoid redundancy of potentially correlated functions, we conducted a Spearman correlation analysis and discarded those with a correlation exceeding 70% (Figure 24). We selected all the functions that we would consider in the study and that are part of the multifunctionality index upon which we will base our analyses.

We selected a total of 23 soil functions: nutrient stocks (TOC, total available N, and available P), enzymatic rates (xylanase, phosphatase, β-N-acetilglucosamininidase, β-glucosidase), N cycle processes (N transformation rate and depolymerization), infiltration, water holding capacity, bulk density, leaf damage resistance (caused by fungi and herbivores), mycorrhizae, root eating nematodes, soil erosion resistance, ecosystem production stability and N and P fertilizer retention capacity. We included methanotroph abundance as a proxy of the regulation of the soil C cycle and climate due to the ability to consume methane of these microorganisms. Root-eating nematodes were also considered a soil function due to their role as pests in croplands.

After standardizing the values of each function (using Z-scores, at the country level), we used the averaging approach (Maestre et al. 2012) as a standard and easy to interpret methodology to obtain a multifunctionality metric. This is calculated simply as the average of the standardized values of all functions measured within a given site.

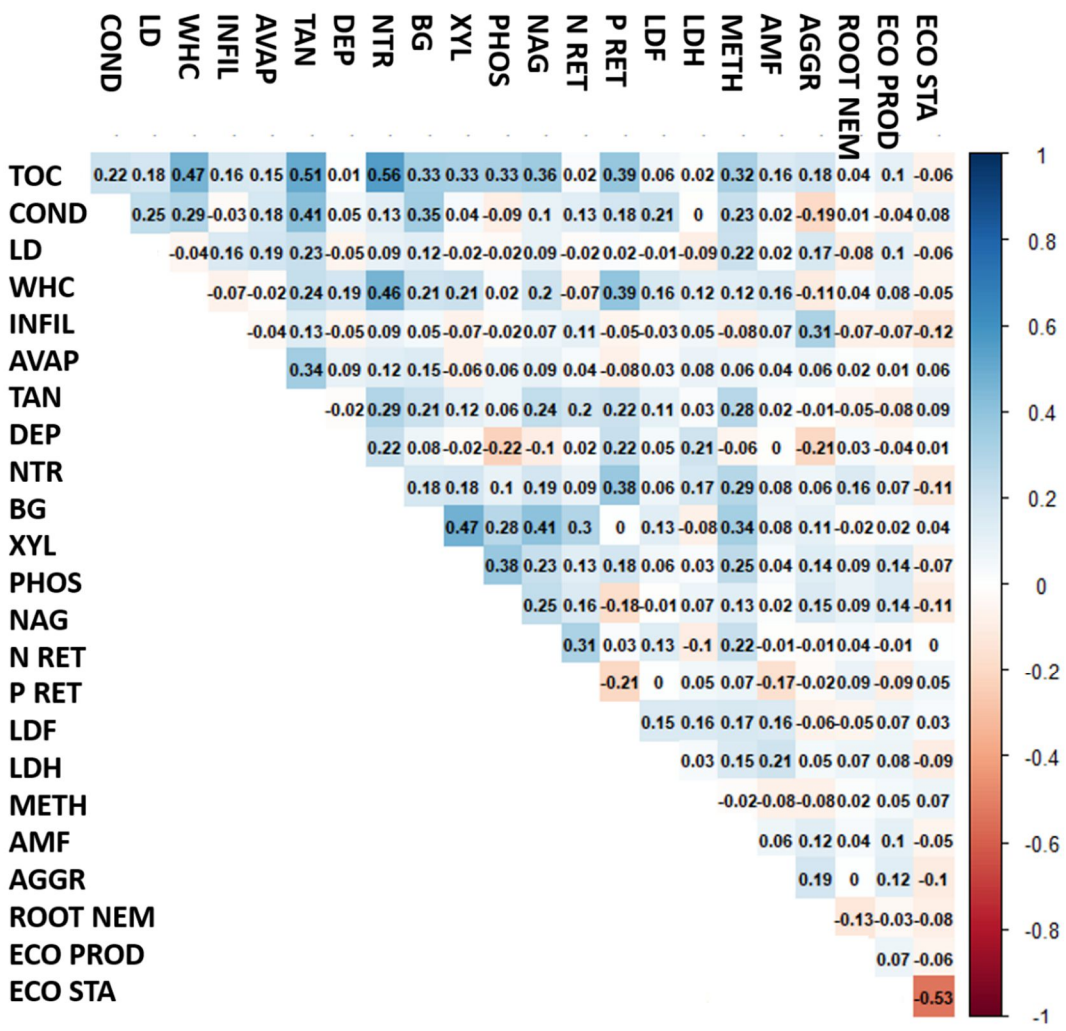


Figure 24. *Correlation matrix (Spearman correlation coefficient) among soil functions measured in the study for EU NUTS-2 and international regions. Negative and positive relationships between each pair of variables are represented in blue and red hues, respectively. Abbreviations → TOC: Total organic carbon; CONDUC: Conductivity; LD: Litter decomposition; WHC: Water holding capacity; INFILTR: Infiltration; AVAP: Available P; TAN: Total available N; DEP: Potential depolymerization rate; NTR: Potential N transformation rate; BG: Activity of b-glucosidase; XYL: Activity of Xylanase; PHOS: Activity of phosphatase ; NAG: Activity of b-N-acetylglucosaminidase; N RET: N retained by soil (lixiviates); P RET: P retained by soil (lixiviates); LDF: Leaf damage fungi; LDH: Leaf damage herbivores; METH: Methanotrophs; AMF: Arbuscular mycorrizal fungi; AGGR: Aggregates (soil erosion resistance); ROOT NEM: Root nematodes abundance; ECO PROD: Ecosystem production; ECO STA: Ecosystem stability.*

4.3. Relative impact of soil management in soil multifunctionality

Similar to the biodiversity groups (section 3 above), we assessed the relative impact of soil management practices on the soil functions quantified for each cropland site of the cross-biome network of sites. These analyses considered 23 functioning indicators (grouped by its relation to NCP: food production, soil formation and protection, climate regulation, regulation of hazards and extreme events, regulation of detrimental organisms and regulation of freshwater quality) in response to agricultural management. Analyses for the grasslands and forests sites of Ireland and Finland were conducted independently due to the differing characteristics of vegetation, soil, and soil management.

We observed general positive responses to alternative soil management in soil organic carbon, the rate of organic matter decomposition, and soil enzymatic activity (phosphatase and xylanase activity) (Figure 25). These findings suggest that alternative soil management could promote higher rates of soil

nutrient recirculation, thereby enhancing their availability for primary production(Peltoniemi et al., 2021; Rani et al., 2023; Sofo et al., 2022).

Conversely, conventional soil management favors an increase in the rate of N transformation. In conventional agrosystems, the addition of mineral nitrogen fertilizers would mean a lower C:N ratio and therefore higher rates of N transformation such as nitrification or ammonification (Bengtsson et al., 2003). However, alternative soil management could favor a higher rate of N immobilization by microorganisms due to an increase in the carbon-to-nitrogen ratio (Cao et al., 2021; Plante and Parton, 2007).

Furthermore, both methanotrophs abundance and root-eating nematodes is favored by alternative soil management, which could improve the balance of greenhouse gas exchanges in the soil acting as a climate regulators (Guerra et al., 2021; Tiwari et al., 2018) but could affect negatively crop productivity by enhancing root damage (but see Jaffuel et al., 2016; Koppenhöfer et al., 2020). Our findings highlight that alternatively managed agrosystems promotes the increase of soil organic carbon and nutrient recycling processes and stocks although they may sustain higher levels of damage by potential pests.

CROPLANDS. ALTERNATIVE VS CONVENTIONAL

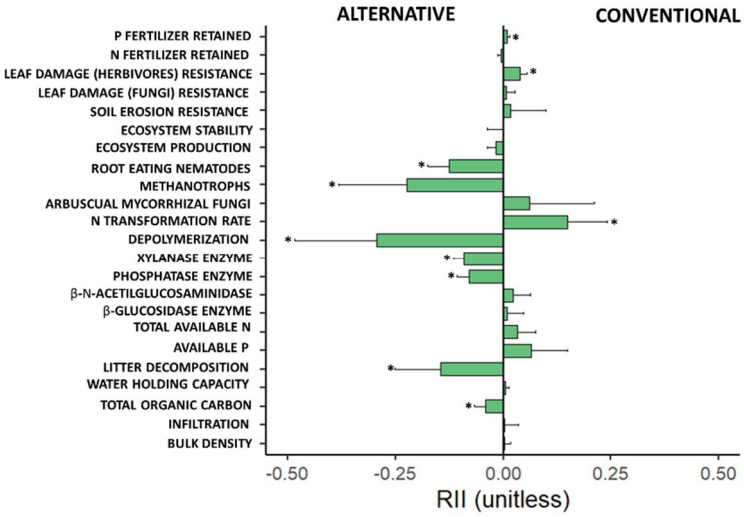
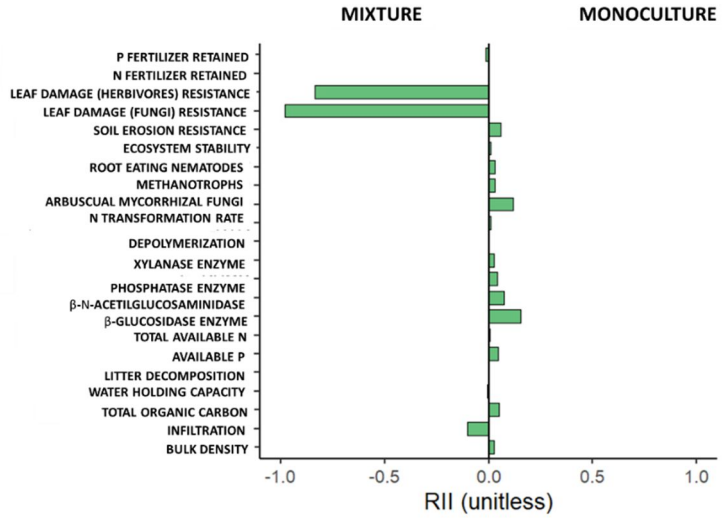


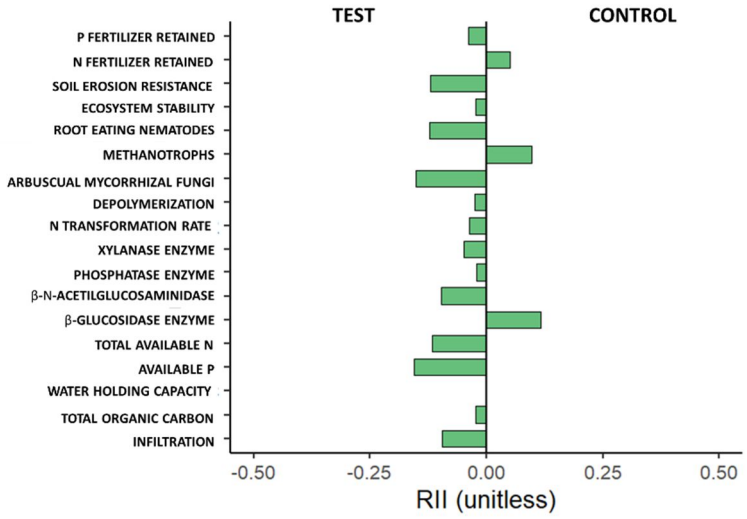
Figure 25. *Relative interaction index (RII) for the 23 soil functions variables measured in croplands for 5 EU NUTS-2 and 3 international regions. RII was calculated for each of the 10 pairs of conventional vs alternative, irrespective of their soil degradation level. Negative RIIs show higher biodiversity levels in alternative vs conventional agriculture, whereas positive RIIs show the contrary. Asterisks indicate for which organisms these differences are significant (p value <0.05).*

We also found consistency in the effect of sustainable agricultural soil management practices in Finland and agroforestry (test) in Ireland for most soil functions (Figure 26). In Ireland, contrary to expectations, we only found a positive relative impact of mixed plant species cover for resistance to leaf damage caused by fungi or pathogens. However, these results lack statistical support because the RII calculation could only be performed independently for the plots belonging to either the Irish agroecosystems or the boreal forests.

IRELAND. MIXTURE VEGETATION VS MONOCULTURE



IRELAND. TEST VS CONTROL (AGROFORESTRY)



FINLAND: CONTINUOUS COVER VS CLEAR CUT FORESTRY

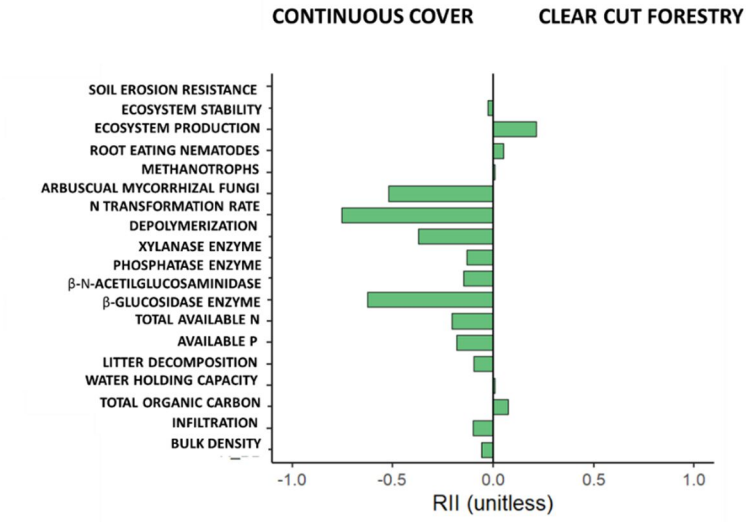


Figure 26. *Relative interaction index (RII) for the soil functions variables measured in grasslands and forests of Ireland (above and middle) and Finland (below). RII was calculated for each of the 10 pairs of mixture vs monoculture and test vs control for Ireland and continuous cover vs clear cut forestry for Finland. Negative RIIs show higher functions levels in mixture and continuous cover agrosystems, whereas positive RIIs show higher functions levels in monoculture and clear-cut forestry.*

4.4. Impact of soil degradation, soil management and soil biodiversity on multifunctionality

4.4.1. Soil biodiversity potential to deliver soil multifunctionality

Many studies show that biodiversity is one of the factors that most determines multifunctionality in terrestrial ecosystems (Isbell et al., 2015; Lefcheck et al., 2015; Meyer et al., 2018). Hence, in addition to fundamental soil attributes, landscape context, and climate, it is imperative to consider the microbial communities and soil mesofauna associated with the soil, along with their interactions with agricultural practices, aridity, or soil degradation.

To consider all these effects, we produced linear models with all soil functions and also for a general metric of "multifunctionality" including as predictors: soil management x soil degradation x aridity x biodiversity + % naturalness + surface hedgerows + sand content + soil pH. This model was simplified using the stepAIC function to obtain the most parsimonious model for each soil function. We did this for both EU NUTS-2 regions (Table S1 in Annex IV) and the three international regions (Table S2 in Annex IV) as a complementary sensitivity analysis.

Our results indicate that **both soil management and biodiversity are key predictive factors of multifunctionality in the European croplands** (Tables S1 and 2, figure 27 and 28). Both factors exerted a significant influence on soil organic carbon concentration and xylanase enzymatic activity, with higher values found under an alternative soil management strategy. Additionally, we observed that biodiversity (i.e. biodiversity index) also affects soil nitrogen transformation rates. In contrast, agricultural soil management also emerged as a significant predictor for both soil aggregate stability (an indicator of erosion resistance) and litter decomposition rate, showing consistent increases in both cases under alternative soil. These findings were consistent across all regions in the joint analysis, although in some cases with lower statistical robustness.

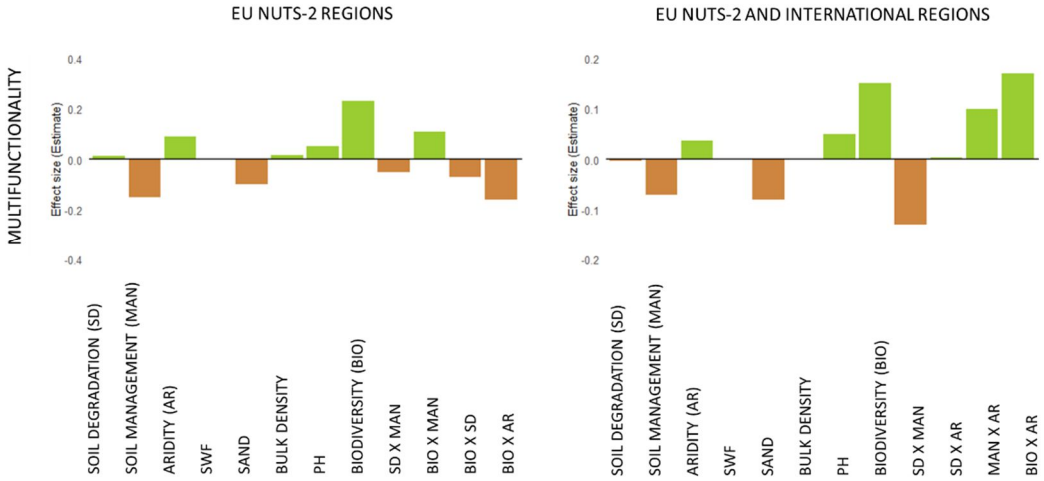


Figure 27. *Effect size of the different soil physico-chemical properties, landscapes attributes, soil agricultural management and degradation state and their interactions on the multifunctionality for European (left side) NUTS-2 and European NUTS-2 and international (right side) regions. Effect sizes (t-values) of the linear models are represented in green and brown for positive and negative effects, respectively. SWF: small wood features.*

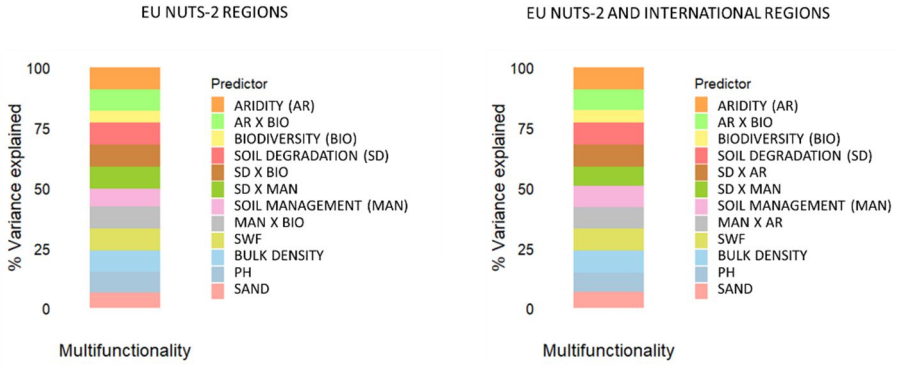


Figure 28. *Variance partitioning illustrating the relative importance of different soil physico-chemical properties, landscapes attributes, soil agricultural management and degradation state and their interactions on multifunctionality for European NUTS-2 (left side) and European NUTS-2 and international regions (right side). NEI: naturalness evaluation index; SWF: small wood features.*

4.4.2. Evaluating the consequences and potential trade-offs to consider when transitioning from conventional to organic agriculture

Despite the benefits of alternative agriculture on soil biodiversity and multifunctionality, it may present controversies due to lower production yields and the need for larger arable land compared to conventional agriculture (Muller et al., 2017; Ponisio et al., 2015; Seufert et al., 2012). Muller et al. (2017) demonstrated, in an analysis of various potential scenarios regarding the increase in the percentage of organic farming, that a complete transition to organic production leads to additional soil use (ranging from 16% to 33%), exacerbated further when considering the adverse effects of climate change on yields. Other factors, such as deforestation or nitrogen limitations, would also contribute to the increase in land areas (Ponisio et al., 2015). Thus, we investigated the transition commitments to organic agriculture regarding biodiversity, multifunctionality, crop yield, and losses or nitrogen dependency of agroecosystems.

To test the different outcomes between conventional and alternative soil management from the most applied perspective possible, we considered: multifunctionality, biodiversity, crop yield, N limitation (i.e. total available N) and N losses (N fertilizer retained and denitrifying genes). These have major implications for food provision and economical sustainability (crop yield), but also considers one of the

main aims of the EU´s Farm-to-fork strategy, which is halt soil degradation, substantially reduce N loss and protect soil biodiversity. Finally, the selection of variables touches upon one of the major risks to be considered when transitioning from conventional to alternative soil management, which is the potential of N to be strongly limiting. To assess all these potential risks and trade-offs, we performed a Spearman correlation analysis both globally and distinguishing by types of soil management (conventional versus alternative). We assessed correlation analysis and plots using "dplr" and "ggplot2" packages (Wickham et al., 2019).

The transition to organic agriculture does not involve compromises for crop yields and N limitation. Our analysis showed a positive correlation between biodiversity and crop yield under alternative soil management (ρ = 0.16) (Figure 29A). Likewise, multifunctionality and crop yield were positively correlated for both conventional and alternative soil management (ρ = 0.28 and ρ = 0.19, respectively) (Figure 29B). Also, we observed a positive relationship between biodiversity and multifunctionality in both types of agricultural soil management (ρ = 0.18 for conventional and ρ = 0.35 for alternative (Figure 29C). These findings reveal that biodiversity is associated with more fertile soils and consequently higher crop production in alternative soil management. Therefore, it is crucial to adopt agricultural practices that protect and conserve soil biodiversity to ensure optimal yields without compromising the economic viability of agroecosystems.

It is important to note, however, that we found some conventional sites showing soil biodiversity levels as high as those observed under alternative soil management. **This suggests a clear potential for refining management recommendations towards enhancing sustainable practices, regardless of the conventional vs alternative (or organic) dichotomy**. According to current literature, sustainable practices that could maximize environmental benefits while maintaining conventional soil management include reduced tillage, organic amendments, cover crops and crop rotation.

Further, we found positive relationships between total available nitrogen and crop yield (ρ = 0.32 and 0.26 for conventional and alternative soil management, respectively) (Figure 29D). **These results reveal that conventional agroecosystems are more dependent than alternative ones on N mineral fertilizers for increasing food production**. Several meta-analyses indicate a yield increase of +6% in conventional agriculture compared to conservation (i.e. organic) agriculture and +19 to +25% compared to alternative agriculture (Knapp and van der Heijden, 2018; Seufert et al., 2012; Wittwer et al., 2021), which would indicate a larger proportion of land extension of organic farming (Muller et al., 2017). However, this approach may lead to lower multifunctionality and higher indirect environmental costs, such as soil degradation and aquifer contamination, due to the reliance on fertilizers for crop production in conventional agroecosystems.

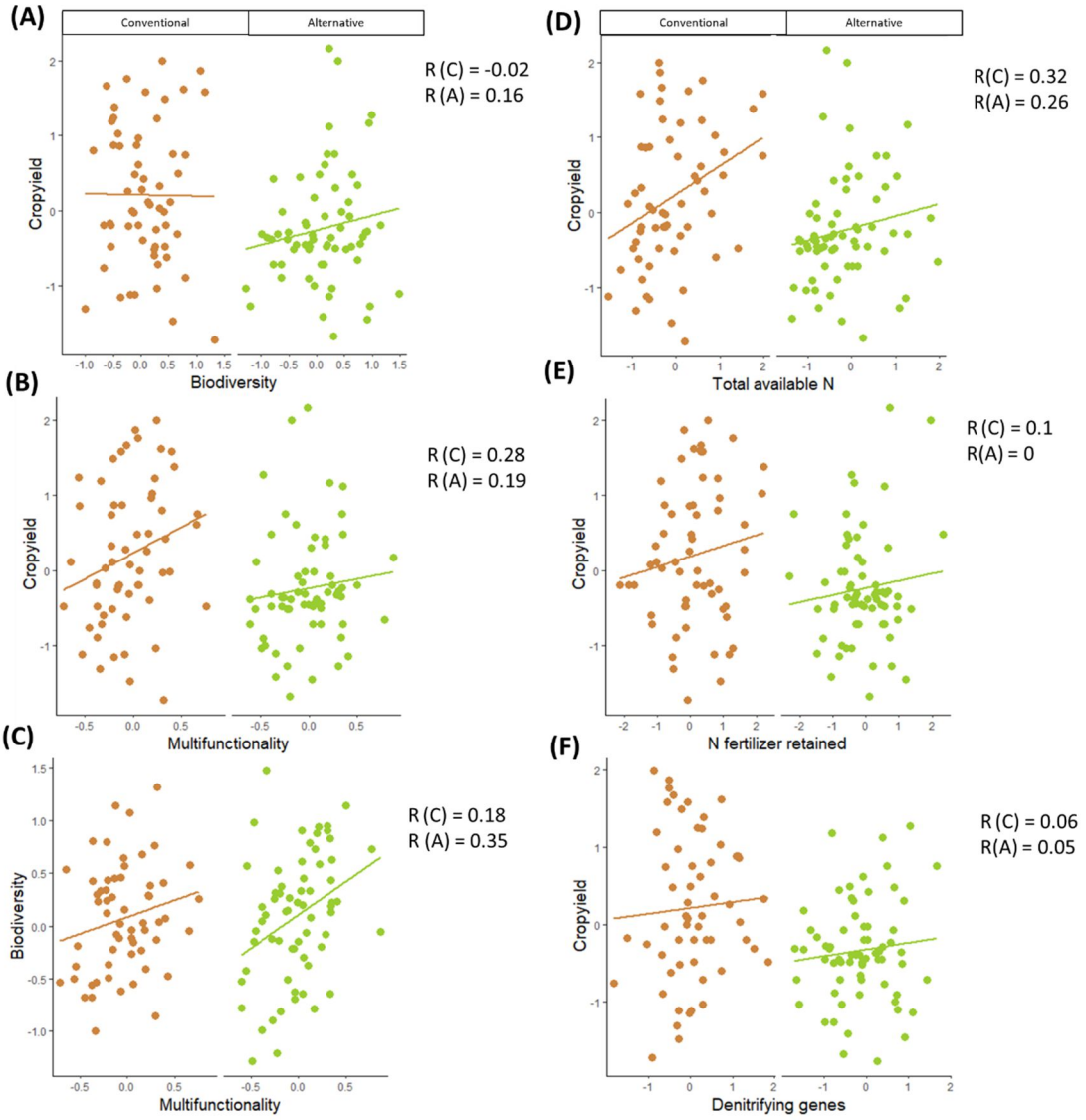


Figure 29. *Relationship (Spearman's correlations [R]) between biodiversity (i.e. biodiversity index), multifunctionality, crop yield, N fertilizer retained, total available N and denitrifying genes differentiating between conventional (C) and alternative soil management (A) (brown and green circles, respectively). All variables are standardized by z-score (unitless).*

The European Commission has set the objective within its "Farm to Fork" strategy to increase the proportion of agricultural land under organic soil management from 9.1% in 2020 (Eurostat, 2023) to 25% by 2030 (European Commission, 2021). But will this be sufficient to halt soil biodiversity loss? What can we expect from this increase in organic agriculture in terms of soil biodiversity? To address these questions, we simulated 500 5-site landscapes in a statistical randomization analysis using all our cropland sites, and determined the optimal percentage of alternative and conventional agriculture for maximizing biodiversity.

To be able to analyze them collectively, we use region-based standardizations using z-scores, such as (x-mean)/st.deviation. Where x = the value of each variable in a particular site, mean = mean of all values of that variable in all sites within that region, and st.deviation = standard deviation of all values of that variable in all sites within that region.

We selected the simulated landscapes that maximized soil biodiversity (standardized average across abundance or diversity levels of all organisms measured) (Allan et al., 2015). We then quantified

the percentage of conventional vs alternative agriculture in those simulated landscapes exhibiting maximum diversity levels. Additionally, we categorised our sites based on low, medium, and high soil degradation levels to evaluate whether the percentage of required alternative agriculture to maximize landscape diversity changed in response to this factor. Futhermore, we identified the simulated landscapes that maximized soil multifunctionality (standardized average of soil functions) and crop yield (kg per hectare of cultivation). All analyses were performed using the R package “vegan” (Oksanen et al., 2020) and “reshape2” (Wickham, 2007) in R version 4.1 (R Core Team, 2020).

We selected the simulated landscapes that maximized soil biodiversity, multifunctionality and crop yield (see Lopezosa et al. 2024 for a related approach). The simulations of **landscape mixtures that maximized all three agrosystems attributes -soil biodiversity, ecosystem multifunctionality and crop yield- indicated that >50 % of the agriculture** in these landscapes was alternative (Figure 32). **This proportion remained consistent regardless of the level of soil degradation** (~48% and ~53% in highly vs low degraded soils, respectively (Figure 32 B). In highly degraded soils, the percentage of alternative soil management in optimal landscapes was ~38-49% (crop yield vs multifunctionality), while it ranged between ~35-61% in soils moderately degraded (crop yield vs biodiversity) (Figures S1-B, S2-B and S3-B in Annex IV).

To maximize biodiversity, we observed a higher percentage of conventional soil management (~60%) in highly degraded soils (Figure S1-B in Annex IV). This increase in the proportion of conventional soil management might be due to an increase in the diversity of "stress-tolerant bacteria, which are known for their resilience to intense soil uses (De Vries and Shade, 2013; Delgado-Baquerizo et al., 2017). Moreover, this is further supported by the positive trend in the relative importance of prokaryotes in conventional soil management for highly degraded soils.

These findings strongly suggest that the **European Union's goal (Farm to Fork strategy) of achieving 25% organic agriculture for 2030 is is insufficient for ensuring optimal levels of soil biodiversity and multifunctionality**. According to our results, the EU would require as much as 50% alternative (i.e. organic) agriculture to maximize biodiversity, multifunctionality and crop yield. These pillars should be considered to provide a comprehensive view of the proportion of alternative and conventional agriculture necessary to ensure crop viability from both ecological and economic perspectives.Enhancing soil quality (increases in organic carbon storage and improving soil fertility) may decrease the likelihood of soil degradation across physical, chemical, biological, and ecological dimensions (Gaikwad et al., 2023). Consequently, efforts to transition from conventional to alternative farming should prioritize areas with medium soil degradation levels. This is emphasized by the fact that a larger proportion of alternative agriculture is necessary to maximize landscape diversity in areas with medium rather than high levels of soil degradation.

Thus, establishing an optimized proportion between cropping systems (i.e., conventional and alternative soil management) could be a strategy to simultaneously achieve a balance between satisfactory yields and environmental integrity, improving soil biodiversity and multifunctionality. Therefore, this approach is highly relevant for agricultural soil management policies in Europe, as it promotes informed decision-making in agrosystem soil management and ensures sustainability in the medium and long term.

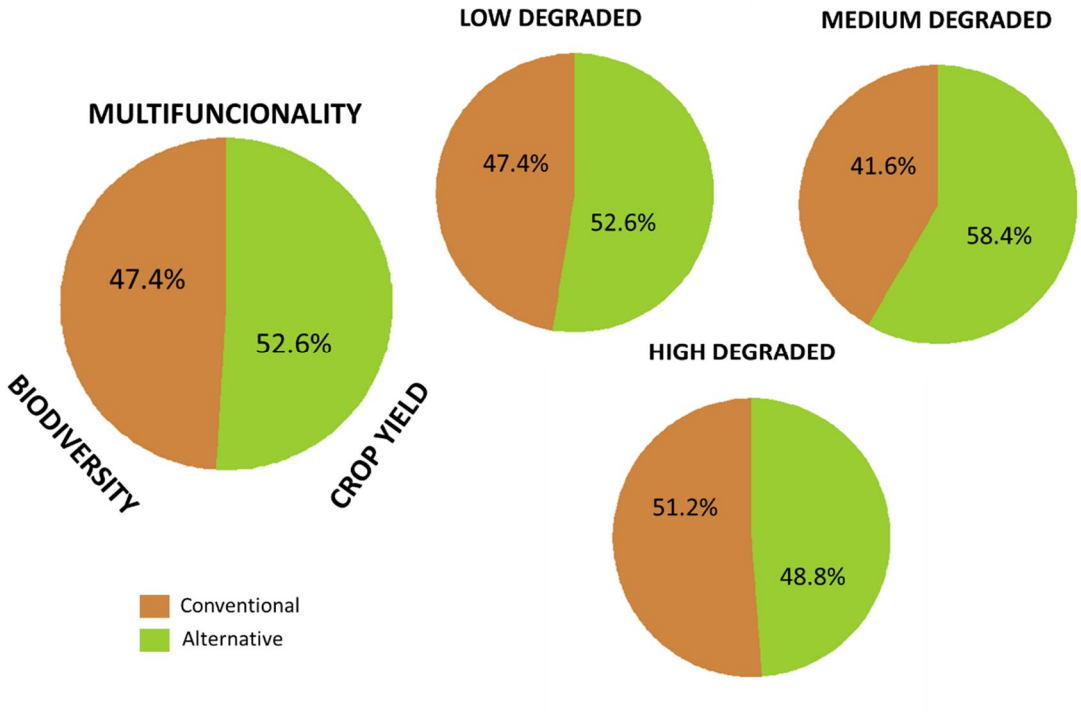


***Figure 30**. Simulated landscapes (using observed field data) maximizing multifunctionality, soil biodiversity and crop yield (A) at the landscape (5 pooled sites) scale and considering soil degradation levels (medium and high degraded) (B). Different colours show the proportion of conventional (brown) vs alternative (green) soil management required to maximize soil biodiversity across all our sites. To obtain the highest values of biodiversity across organisms and sites, we used a "biodiversity index" averaging the standardized values for each soil organism considered.*

5. Conclusions

Although still in the preliminary stage, our search for reliable and monetary cost-effective biodiversity indicators suggests that combining total microbial biomass with the quantification of a target gene related to the nitrogen cycle (such as aob, oa, coma_a, or coma_b) could serve as a valuable set of soil biodiversity bioindicators. . Given the significant mismatch we observed between molecular and taxonomic approaches when examining nematodes and microarthropods, along with the contrasting environmental responses of the latter, incorporating the abundance of collembolans or acari into this indicator list could provide a more comprehensive evaluation.

Despite strong idiosyncracies depending on site, taxa, or function of interest, we found clear benefits of alternative agriculture, both in terms of soil biodiversity (higher alpha- and beta-diversity, more connected “brown webs”), and increasing soil organic carbon and nutrient cycling. Equally important is the lack of strong negative effects associated with this shift from conventional to alternative practices; generally, biodiversity or functioning indicators that did not respond positivelly, showed no significant changes in response to soil management.

We did not find evidence of strong trade-offs between the main axes of the EU farm-to-fork strategy, as we observed either positive or neutral relationships, but no negative relationships, between soil biodiversity, crop yield, N loss, or soil multifunctionality. The absence of significant trade-offs suggests that the EU's ambitions of establishing 25% organic agriculture by 2030 may be insufficient. Indeed, the findings presented in this report indicate that raising this target to 50% would facilitate the simultaneous maximisation of biodiversity, multifunctionality, and crop yield. Moreover, this conversion would likely be more effective in areas with moderately to highly degraded soils. Our results also indicate a strong potential to achieve high levels of soil biodiversity and health, even within conventional agricultural systems.

6. References

Allan, E., Bossdorf, O., Dormann, C.F., Prati, D., Gossner, M.M., Tscharntke, T., Blüthgen, N., Bellach, M., Birkhofer, K., Boch, S., Böhm, S., Börschig, C., Chatzinotas, A., Christ, S., Daniel, R., Diekötter, T., Fischer, C., Friedl, T., Glaser, K., Hallmann, C., Hodac, L., Hölzel, N., Jung, K., Klein, A.M., Klaus, V.H., Kleinebecker, T., Krauss, J., Lange, M., Morris, E.K., Müller, J., Nacke, H., Pašalić, E., Rillig, M.C., Rothenwöhrer, C., Schall, P., Scherber, C., Schulze, W., Socher, S.A., Steckel, J., Steffan-Dewenter, I., Türke, M., Weiner, C.N., Werner, M., Westphal, C., Wolters, V., Wubet, T., Gockel, S., Gorke, M., Hemp, A., Renner, S.C., Schöning, I., Pfeiffer, S., König-Ries, B., Buscot, F., Linsenmair, K.E., Schulze, E.D., Weisser, W.W., Fischer, M., 2014. Interannual variation in land-use intensity enhances grassland multidiversity. Proceedings of the National Academy of Sciences of the United States of America 111, 308–313. doi:10.1073/pnas.1312213111

Allan, E., Manning, P., Alt, F., Binkenstein, J., Blaser, S., Blüthgen, N., Böhm, S., Grassein, F., Hölzel, N., Klaus, V.H., Kleinebecker, T., Morris, E.K., Oelmann, Y., Prati, D., Renner, S.C., Rillig, M.C., Schaefer, M., Schloter, M., Schmitt, B., Schöning, I., Schrumpf, M., Solly, E., Sorkau, E., Steckel, J., Steffen-Dewenter, I., Stempfhuber, B., Tschapka, M., Weiner, C.N., Weisser, W.W., Werner, M., Westphal, C., Wilcke, W., Fischer, M., 2015. Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition. Ecology Letters 18, 834–843. doi:10.1111/ele.12469

Anderson, M.J., Willis, T.J., 2003. Canonical analysis of principal coordinates: A useful method of constrained ordination for ecology. Ecology 84, 511–525. doi:10.1890/0012-9658(2003)084[0511:CAOPCA]2.0.CO;2

Armas, C., Ordiales, R., Pugnaire, F.I., 2004. Measuring plant interactions: A new comparative index. Ecology 85, 2682–2686. doi:10.1890/03-0650

Bainard, L.D., Koch, A.M., Gordon, A.M., Klironomos, J.N., 2012. Temporal and compositional differences of arbuscular mycorrhizal fungal communities in conventional monocropping and tree-based intercropping systems. Soil Biology and Biochemistry 45, 172–180. doi:10.1016/j.soilbio.2011.10.008

Barral, M.P., Rey Benayas, J.M., Meli, P., Maceira, N.O., 2015. Quantifying the impacts of ecological restoration on biodiversity and ecosystem services in agroecosystems: A global meta-analysis. Agriculture, Ecosystems and Environment 202, 223–231. doi:10.1016/j.agee.2015.01.009

Benckiser, G. (Ed. ., 1997. Fauna in soil ecosystems: recycling processes, nutrient fluxes, and agricultural production. CRC Press.

Bengtsson-Palme, J., Hartmann, M., Eriksson, K.M., Pal, C., Thorell, K., Larsson, D.G.J., Nilsson, R.H., 2015. metaxa2: Improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. Molecular Ecology Resources 15, 1403–1414. doi:10.1111/1755-0998.12399

Bengtsson, G., Bengtson, P., Månsson, K.F., 2003. Gross nitrogen mineralization-, immobilization-, and nitrification rates as a function of soil C/N ratio and microbial activity. Soil Biology and Biochemistry 35, 143–154. doi:10.1016/S0038-0717(02)00248-1

Betancur-Corredor, B., Lang, B., Russell, D.J., 2022. Reducing tillage intensity benefits the soil micro- and mesofauna in a global meta-analysis. European Journal of Soil Science 73, 1–15. doi:10.1111/ejss.13321

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.J., Taberlet, P., 2012. Tracking earthworm communities from soil DNA. Molecular Ecology 21, 2017–2030. doi:10.1111/j.1365-294X.2011.05407.x

Blake, G.R., Hartge, K.H., 1986. Bulk density. Methods of Soil Analysis, Part 1: Physical and Mineralogical Methods 9, 363–375. doi:10.2136/sssabookser5.1.2ed.c13

Bongers, T., 1990. The maturity index: an ecological measure of environmental disturbance based on nematode species composition. Oecologia 83, 14–19. doi:10.1007/BF00324627

Cao, Y., He, Z., Zhu, T., Zhao, F., 2021. Organic-C quality as a key driver of microbial nitrogen immobilization in soil: A meta-analysis. Geoderma 383, 114784. doi:10.1016/j.geoderma.2020.114784

Chang, L., Wu, H., Wu, D., Sun, X., 2013. Effect of tillage and farming management on Collembola in marsh soils. Applied Soil Ecology 64, 112–117. doi:10.1016/j.apsoil.2012.11.007

Cho, S.J., Kim, M.H., Lee, Y.O., 2016. Effect of pH on soil bacterial diversity. Journal of Ecology and Environment

40, 1–9. doi:10.1186/s41610-016-0004-1

De Vries, F.T., Shade, A., 2013. Controls on soil microbial community stability under climate change. Frontiers in Microbiology 4, 1–16. doi:10.3389/fmicb.2013.00265

de Vries, F.T., Wallenstein, M.D., 2017. Below-ground connections underlying above-ground food production: a framework for optimising ecological connections in the rhizosphere. Journal of Ecology 105, 913–920. doi:10.1111/1365-2745.12783

Delgado-Baquerizo, M., Eldridge, D.J., Ochoa, V., Gozalo, B., Singh, B.K., Maestre, F.T., 2017. Soil microbial communities drive the resistance of ecosystem multifunctionality to global change in drylands across the globe. Ecology Letters 20, 1295–1305. doi:10.1111/ele.12826

Dick, R.P., Dick, L.K., Deng, S., Li, X., Kandeler, E., Poll, C., Freeman, C., Jones, T.G., Weintraub, M.N., Esseili, K.A., Saxena, J., 2018. Cross-laboratory comparison of fluorimetric microplate and colorimetric bench-scale soil enzyme assays. Soil Biology and Biochemistry 121, 240–248. doi:10.1016/j.soilbio.2017.12.020

Du Preez, G., Daneel, M., De Goede, R., Du Toit, M.J., Ferris, H., Fourie, H., Geisen, S., Kakouli-Duarte, T., Korthals, G., Sánchez-Moreno, S., Schmidt, J.H., 2022. Nematode-based indices in soil ecology: Application, utility, and future directions. Soil Biology and Biochemistry 169. doi:10.1016/j.soilbio.2022.108640

Elbrecht, V., Braukmann, T.W.A., Ivanova, N. V., Prosser, S.W.J., Hajibabaei, M., Wright, M., Zakharov, E. V., Hebert, P.D.N., Steinke, D., 2019. Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ 2019, 1–23. doi:10.7717/peerj.7745

European Commission. (2021). Farm to Fork Strategy: For a fair, healthy and environmentally-friendly food system. Retrieved from https://ec.europa.eu/food/farm2fork_en

Eurostat. (2023). Agricultural production - organic farming. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Agricultural_production_-_organic_farming

Fernández, J.A., Niell, F.X., Lucena, J., 1985. A rapid and sensitive automated determination of phosphate in natural waters. Limnology and Oceanography 30, 227–230. doi:10.4319/lo.1985.30.1.0227

Ferris, H., Bongers, T., De Goede, R.G.M., 2001. A framework for soil food web diagnostics: Extension of the nematode faunal analysis concept. Applied Soil Ecology 18, 13–29. doi:10.1016/S0929-1393(01)00152-4

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37, 4302–4315. doi:10.1002/joc.5086

Field, D.J., David C. McKenzie B, Koppi, A.J., 1997. Development of an improved Vertisol stability test for SOILpak. Australian Journal of Soil Research 35. doi:10.1071/SR96118

Frey, B., Rime, T., Phillips, M., Stierli, B., Hajdas, I., Widmer, F., Hartmann, M., 2016. Microbial diversity in European alpine permafrost and active layers. FEMS Microbiology Ecology 92, 1–17. doi:10.1093/femsec/fiw018

Gaikwad, A.S., Phule, M., Vidyapeeth, K., Margal, P.B., Titirmare, N.S., Phule, M., Vidyapeeth, K., 2023. Soil Degradation and Remediation: Strategies for Restoring Soil Quality. doi:10.5281/zenodo.8330919

Geisen, S., Snoek, L.B., ten Hooven, F.C., Duyts, H., Kostenko, O., Bloem, J., Martens, H., Quist, C.W., Helder, J.A., van der Putten, W.H., 2018. Integrating quantitative morphological and qualitative molecular methods to analyse soil nematode community responses to plant range expansion. Methods in Ecology and Evolution 9, 1366–1378. doi:10.1111/2041-210X.12999

George, P.B.L., Keith, A.M., Creer, S., Barrett, G.L., Lebron, I., Emmett, B.A., Robinson, D.A., Jones, D.L., 2017. Evaluation of mesofauna communities as soil quality indicators in a national-level monitoring programme. Soil Biology and Biochemistry 115, 537–546. doi:10.1016/j.soilbio.2017.09.022

Gomiero, T., Pimentel, D., Paoletti, M.G., 2011. Environmental impact of different agricultural management practices: Conventional vs. Organic agriculture. Critical Reviews in Plant Sciences 30, 95–124. doi:10.1080/07352689.2011.554355

Griffiths, B.S., de Groot, G.A., Laros, I., Stone, D., Geisen, S., 2018. The need for standardisation: Exemplified by a description of the diversity, community structure and ecological indices of soil nematodes. Ecological

Indicators 87, 43–46. doi:10.1016/j.ecolind.2017.12.002

Guerra, C.A., Delgado-Baquerizo, M., Duarte, E., Marigliano, O., Görgen, C., Maestre, F.T., Eisenhauer, N., 2021. Global projections of the soil microbiome in the Anthropocene. Global Ecology and Biogeography 30, 987–999. doi:10.1111/geb.13273

Guppy, C.N., Menzies, N.W., Moody, P.W., Compton, B.L., Blamey, F.P.C., 2000. A simplified, sequential, phosphorus fractionation method. Communications in Soil Science and Plant Analysis 31, 1981–1991. doi:10.1080/00103620009370556

Harris, D., Horwáth, W.R., van Kessel, C., 2001. Acid fumigation of soils to remove carbonates prior to total organic carbon or CARBON-13 isotopic analysis. Soil Science Society of America Journal 65, 1853–1856. doi:10.2136/sssaj2001.1853

He, J., Zhou, T., Shen, X., Zhang, N., Sun, C., Lu, S., Shao, Y., 2023. Primer selection impacts the evaluation of microecological patterns in environmental microbiomes. IMeta 2. doi:10.1002/imt2.135

Hess, H.H., Derr, J.E., 1975. Assay of inorganic and organic phosphorus in the 0.1-5 nanomole range. Analytical Biochemistry 63, 607–613. doi:10.1016/0003-2697(75)90388-7

Hole, D.G., Perkins, A.J., Wilson, J.D., Alexander, I.H., Grice, P. V., Evans, A.D., 2005. Does organic farming benefit biodiversity? Biological Conservation 122, 113–130. doi:10.1016/j.biocon.2004.07.018

Hugerth, L.W., Muller, E.E.L., Hu, Y.O.O., Lebrun, L.A.M., Roume, H., Lundin, D., Wilmes, P., Andersson, A.F., 2014. Systematic design of 18S rRNA gene primers for determining eukaryotic diversity in microbial consortia. PLoS ONE 9. doi:10.1371/journal.pone.0095567

Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T.M., Bonin, C., Bruelheide, H., De Luca, E., Ebeling, A., Griffin, J.N., Guo, Q., Hautier, Y., Hector, A., Jentsch, A., Kreyling, J., Lanta, V., Manning, P., Meyer, S.T., Mori, A.S., Naeem, S., Niklaus, P.A., Polley, H.W., Reich, P.B., Roscher, C., Seabloom, E.W., Smith, M.D., Thakur, M.P., Tilman, D., Tracy, B.F., Van Der Putten, W.H., Van Ruijven, J., Weigelt, A., Weisser, W.W., Wilsey, B., Eisenhauer, N., 2015. Biodiversity increases the resistance of ecosystem productivity to climate extremes. Nature 526, 574–577. doi:10.1038/nature15374

Jaffuel, G., Mäder, P., Blanco-Perez, R., Chiriboga, X., Fliessbach, A., Turlings, T.C.J., Campos-Herrera, R., 2016. Prevalence and activity of entomopathogenic nematodes and their antagonists in soils that are subject to different agricultural practices. Agriculture, Ecosystems and Environment 230, 329–340. doi:10.1016/j.agee.2016.06.009

Kettler, T.A., Doran, J.W., Gilbert, T.L., 2001. Simplified Method for Soil Particle-Size Determination to Accompany Soil-Quality Analyses. Soil Science Society of America Journal 65, 849–852. doi:10.2136/sssaj2001.653849x

Keuskamp, J.A., Dingemans, B.J.J., Lehtinen, T., Sarneel, J.M., Hefting, M.M., 2013. Tea Bag Index: A novel approach to collect uniform decomposition data across ecosystems. Methods in Ecology and Evolution 4, 1070–1075. doi:10.1111/2041-210X.12097

Kindt, R., Coe, R., 2005. Tree diversity analysis.

Knapp, S., van der Heijden, M.G.A., 2018. A global meta-analysis of yield stability in organic and conservation agriculture. Nature Communications 9, 1–9. doi:10.1038/s41467-018-05956-1

Koppenhöfer, A.M., Shapiro-Ilan, D.I., Hiltpold, I., 2020. Entomopathogenic Nematodes in Sustainable Food Production. Frontiers in Sustainable Food Systems 4, 1–14. doi:10.3389/fsufs.2020.00125

Lefcheck, J.S., Byrnes, J.E.K., Isbell, F., Gamfeldt, L., Griffin, J.N., Eisenhauer, N., Hensel, M.J.S., Hector, A., Cardinale, B.J., Duffy, J.E., 2015. Biodiversity enhances ecosystem multifunctionality across trophic levels and habitats. Nature Communications 6. doi:10.1038/ncomms7936

Li, S., Deng, Y., Wang, Z., Zhang, Z., Kong, X., Zhou, W., Yi, Y., Qu, Y., 2020. Exploring the accuracy of amplicon-based internal transcribed spacer markers for a fungal community. Molecular Ecology Resources 20, 170–184. doi:10.1111/1755-0998.13097

Lopezosa, P., Soliveres, S., Serra, L., Constán-Nava, S., Berdugo, M., 2024. Land use determines Mediterranean ecosystems' multifunctionality more than plant richness or habitat composition. Journal of Applied Ecology

1–13. doi:10.1111/1365-2664.14568

Martins da Silva, P., Carvalho, F., Dirilgen, T., Stone, D., Creamer, R., Bolger, T., Sousa, J.P., 2016. Traits of collembolan life-form indicate land use types and soil properties across an European transect. Applied Soil Ecology 97, 69–77. doi:10.1016/j.apsoil.2015.07.018

Meyer, S.T., Ptacnik, R., Hillebrand, H., Bessler, H., Buchmann, N., Ebeling, A., Eisenhauer, N., Engels, C., Fischer, M., Halle, S., Klein, A.M., Oelmann, Y., Roscher, C., Rottstock, T., Scherber, C., Scheu, S., Schmid, B., Schulze, E.D., Temperton, V.M., Tscharntke, T., Voigt, W., Weigelt, A., Wilcke, W., Weisser, W.W., 2018. Biodiversity-multifunctionality relationships depend on identity and number of measured functions. Nature Ecology and Evolution 2, 44–49. doi:10.1038/s41559-017-0391-4

Mills, A., Fey, M., Donaldson, J., Todd, S., Theron, L., 2009. Soil infiltrability as a driver of plant cover and species richness in the semi-arid Karoo, South Africa. Plant and Soil 320, 321–332. doi:10.1007/s11104-009-9904-5

Muller, A., Schader, C., El-Hage Scialabba, N., Brüggemann, J., Isensee, A., Erb, K.-H., Smith, P., Klocke, P., Leiber, F., Stolze, M., Niggli, U., 2017. Strategies for feeding the world more sustainably with organic agriculture. Nature Communications 8, 1290. doi:10.1038/s41467-017-01410-w

Ning, Q., Chen, L., Zhang, C., Ma, D., Li, D., Han, X., Cai, Z., Huang, S., Zhang, J., 2021. Saprotrophic fungal communities in arable soils are strongly associated with soil fertility and stoichiometry. Applied Soil Ecology 159, 103843. doi:10.1016/j.apsoil.2020.103843

Oksanen, A.J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.R., Hara, R.B.O., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., 2020. Package ' vegan .'

Olsen, S., Cole, C., Watanabe, F., Dean, L., 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. USDA Circular 939, 1–19.

Orgiazzi, A., Bardgett, R.D., Barrios, E., V., B.-P., Briones, M.J.I., Chotte, J-L., De Deyn, G.B., Eggleton, P., Fierer, N., Fraser, T.K., S., H.J., N.C., J., Jones, A., Kandeler, E., Kaneko, N., P., L., Lemanceau, P., Miko, L., Montanarella, L., Moreira, F.M.S., K.S., R., Scheu, S., Singh, B.K., Six, J., van der Putten, 2016. Global Soil biodiversity atlas. doi:doi:10.2788/799182

Parisi, V., Menta, C., Gardi, C., Jacomini, C., Mozzanica, E., 2005. Microarthropod communities as a tool to assess soil quality and biodiversity: A new approach in Italy. Agriculture, Ecosystems and Environment 105, 323–333. doi:10.1016/j.agee.2004.02.002

Paula, F.S., Tatti, E., Thorn, C., Abram, F., Wilson, J., O'Flaherty, V., 2020. Soil prokaryotic community resilience, fungal colonisation and increased cross-domain co-occurrence in response to a plant-growth enhancing organic amendment. Soil Biology and Biochemistry 149, 107937. doi:10.1016/j.soilbio.2020.107937

Pellegrino, E., Piazza, G., Helgason, T., Ercoli, L., 2021. Eukaryotes in soil aggregates across conservation managements: Major roles of protists, fungi and taxa linkages in soil structuring and C stock. Soil Biology and Biochemistry 163, 108463. doi:10.1016/j.soilbio.2021.108463

Peltoniemi, K., Velmala, S., Fritze, H., Lemola, R., Pennanen, T., 2021. Long-term impacts of organic and conventional farming on the soil microbiome in boreal arable soil. European Journal of Soil Biology 104, 103314. doi:10.1016/j.ejsobi.2021.103314

Ponisio, L.C., M'gonigle, L.K., Mace, K.C., Palomino, J., Valpine, P. De, Kremen, C., 2015. Diversification practices reduce organic to conventional yield gap. Proceedings of the Royal Society B: Biological Sciences 282. doi:10.1098/rspb.2014.1396

R Core Team, 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Rani, M., Kaushik, P., Bhayana, S., Kapoor, S., 2023. Impact of organic farming on soil health and nutritional quality of crops. Journal of the Saudi Society of Agricultural Sciences 22, 560–569. doi:10.1016/j.jssas.2023.07.002

Reeleder, R.D., Miller, J.J., Ball Coelho, B.R., Roy, R.C., 2006. Impacts of tillage, cover crop, and nitrogen on populations of earthworms, microarthropods, and soil fungi in a cultivated fragile soil. Applied Soil Ecology 33, 243–257. doi:10.1016/j.apsoil.2005.10.006

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: A versatile open source tool for metagenomics. PeerJ 2016, 1–22. doi:10.7717/peerj.2584

Rousk, J., Bååth, E., Brookes, P.C., Lauber, C.L., Lozupone, C., Caporaso, J.G., Knight, R., Fierer, N., 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME Journal 4, 1340–1351. doi:10.1038/ismej.2010.58

Rutgers, M., Schouten, A.J., Bloem, J., Van Eekeren, N., De Goede, R.G.M., Jagers Op Akkerhuis, G.A.J.M., Van Der Wal, A., Mulder, C., Brussaard, L., Breure, A.M., 2009. Biological measurements in a nationwide soil monitoring network. European Journal of Soil Science 60, 820–832. doi:10.1111/j.1365-2389.2009.01163.x

Schimel, J.P., Bennet, J., 2004. Nitrogen mineralization: Challenges of a changing paradigm. Ecology 85, 591–602.

Seufert, V., Ramankutty, N., Foley, J.A., 2012. Comparing the yields of organic and conventional agriculture. Nature 485, 229–232. doi:10.1038/nature11069

Siepel, H., 1996. Biodiversity of soil micro arthropods: The filtering of species. Biodiversity and Conservation 5, 251–260. doi:10.1007/BF00055834

Siepel, H., Vogels, J., Bobbink, R., Bijlsma, R.J., Jongejans, E., de Waal, R., Weijters, M., 2018. Continuous and cumulative acidification and N deposition induce P limitation of the micro-arthropod soil fauna of mineral-poor dry heathlands. Soil Biology and Biochemistry 119, 128–134. doi:10.1016/j.soilbio.2018.01.025

Sieriebriennikov, B., Ferris, H., de Goede, R.G.M., 2014. NINJA: An automated calculation system for nematode-based biological monitoring. European Journal of Soil Biology 61, 90–93. doi:10.1016/j.ejsobi.2014.02.004

Sikder, M.M., Vestergård, M., Sapkota, R., Kyndt, T., Nicolaisen, M., 2020. Evaluation of metabarcoding primers for analysis of soil nematode communities. Diversity 12, 1–14. doi:10.3390/d12100388

Sims, G.K., Ellsworth, T.R., Mulvaney, R.L., 1995. Microscale determination of inorganic nitrogen in water and soil extracts. Communications in Soil Science and Plant Analysis 26, 303–316. doi:10.1080/00103629509369298

Sofo, A., Zanella, A., Ponge, J.F., 2022. Soil quality and fertility in sustainable agriculture, with a contribution to the biological classification of agricultural soils. Soil Use and Management 38, 1085–1112. doi:10.1111/sum.12702

SOILGUARD Team, 2023. Cross-biome network of sites set up (Deliverable 2.1).

SOILGUARD Team, 2023). Soil biodiversity status in European and international biogeographical regions (Deliverable 2.2).

Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Biologiske Shrifter 4, 1–34.

Tedersoo, L., Lindahl, B., 2016. Fungal identification biases in microbiome projects. Environmental Microbiology Reports 8, 774–779. doi:10.1111/1758-2229.12438

Tiwari, S., Singh, C., Singh, J.S., 2018. Land use changes: a key ecological driver regulating methanotrophs abundance in upland soils. Energy, Ecology and Environment 3, 355–371. doi:10.1007/s40974-018-0103-1

Tongway, D., Hindley, N., 2004. LANDSCAPE FUNCTION ANALYSIS : With special reference to Minesites and Rangelands. Cisro Sustainable Ecosystems 2–80.

Trabucco, A., Zomer, R.J., 2018. Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2. CGIAR Consortium for Spatial Information (CGIAR-CSI) 10.

Tuck, S.L., Winqvist, C., Mota, F., Ahnström, J., Turnbull, L.A., Bengtsson, J., 2014. Land-use intensity and the effects of organic farming on biodiversity: A hierarchical meta-analysis. Journal of Applied Ecology 51, 746–755. doi:10.1111/1365-2664.12219

Wang, L., Xin, J., Nai, H., Zheng, X., 2021. Effects of different fertilizer applications on nitrogen leaching losses and the response in soil microbial community structure. Environmental Technology and Innovation 23, 101608. doi:10.1016/j.eti.2021.101608

Wang, W., Yang, Y., Li, J., Bu, P., Lu, A., Wang, H., He, W., Santos Bermudez, R., Feng, J., 2024. Consecutive

Fertilization-Promoted Soil Nutrient Availability and Altered Rhizosphere Bacterial and Bulk Fungal Community Composition. Forests 15, 1–18. doi:10.3390/f15030514

Wickham, H., 2007. Reshaping data with the reshape package. Journal of Statistical Software 21, 1–20. doi:10.18637/jss.v021.i12

Wickham, H., François, R., Henry, L., Müller, K., 2019. dplyr: A Grammar of Data Manipulation. R package version, Media.

Wiesel, L., Daniell, T.J., King, D., Neilson, R., 2015. Determination of the optimal soil sample size to accurately characterise nematode communities in soil. Soil Biology and Biochemistry 80, 89–91. doi:10.1016/j.soilbio.2014.09.026

Wittwer, R.A., Bender, S.F., Hartman, K., Hydbom, S., Lima, R.A.A., Loaiza, V., Nemecek, T., Oehl, F., Olsson, P.A., Petchey, O., Prechsl, U.E., Schlaeppi, K., Scholten, T., Seitz, S., Six, J., Van Der Heijden, M.G.A., 2021. Organic and conservation agriculture promote ecosystem multifunctionality. Science Advances 7, 1–13. doi:10.1126/sciadv.abg6995

Xia, Q., Rufty, T., Shi, W., 2020. Soil microbial diversity and composition: Links to soil texture and associated properties. Soil Biology and Biochemistry 149, 107953. doi:10.1016/j.soilbio.2020.107953

7. Supplementary materials

Annex I. Naturalness Evaluation Index based on the Corine Land Cover dataset across the European NUTS-2 regions

1. Introducing the CLC 2018 dataset & the Naturalness Evaluation Index

The first update of the initial CLC dataset (1989-1990) dates back to 2000 while further datasets/updates followed with an update cycle of 6 years [1]. The CLC 2018 dataset is the newest available CLC product, and thus it was selected for the needs of the project. The dataset is constructed from several categories that describe artificial surfaces, agricultural areas, semi natural areas, forests (high natural areas), wetlands and water bodies (Figure 1). These categories were later used as described in Baiamonte G. et al, 2015 [2] for the calculation of the NEI.

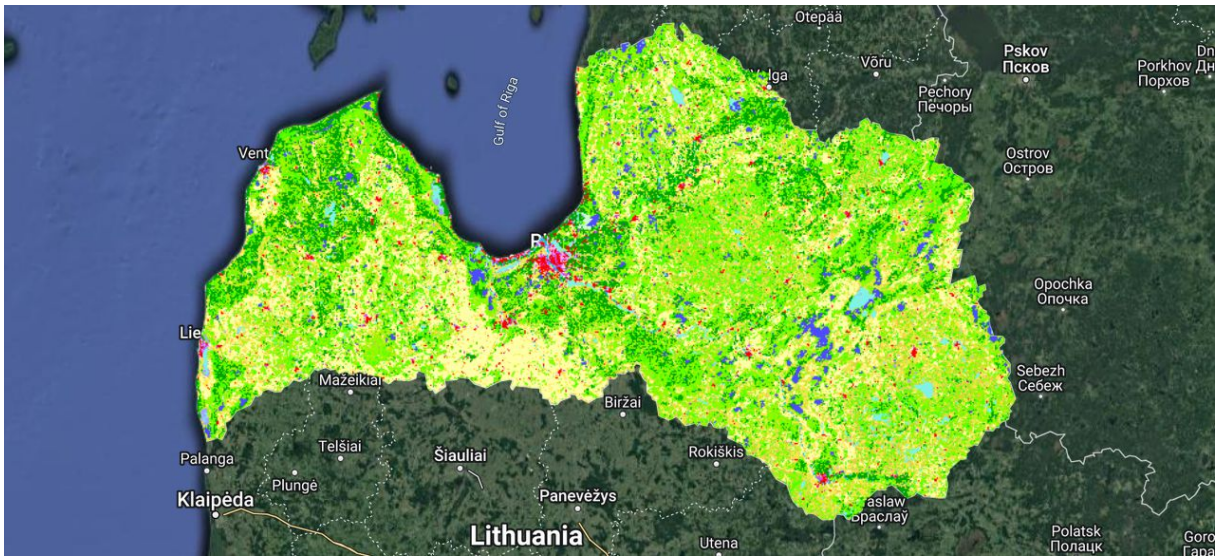


Figure 1. Corine Land Cover 2018 dataset's example NUTS-2 region (Latvia) via Google Earth Engine.

2. CLC 2018 dataset & the Naturalness Evaluation Index

Materials & Methods

For acquiring information about the naturalness of the project's NUTS-2 regions, the CLC 2018 dataset was downloaded by the Google Earth Engine API (Table 1).

Table 1. Materials used and relevant information.

Dataset	Coverage	Spatial Resolution (m)	Temporal Reference	Source / Manual
Corine Land Cover 2018	Pan - European	100	2017 - 2018	GEE / CLC 2018 Manual

The downloaded datasets were visually inspected, manipulated and processed into the ArcGIS Pro environment by using several geospatial tools such as project, clip, intersect, buffer, reclassify, raster to polygon and summary statistics by following the steps of the Baiamonte G. et al, 2015 [2] paper.

In particular, once the datasets were downloaded, they were preprocessed and then the CLC 2018 dataset was reclassified into four classes that resulted in the production of the naturalness map.

These classes refer to:

1. High Naturalness Systems
2. Semi-natural Systems
3. Agricultural Systems and
4. Artificial Systems.

Once the naturalness map was ready, the Naturalness Evaluation Index was calculated for each single pixel as described in the paper and by using the following formula:

$NEI = C_1 + 2C_2 + 3C_3 / 3(C_0 + C_1 + C_2 + C_3)$

Where,

C_0 the area covered by Artificial systems,

C_1 is the area covered by Agricultural systems,

C_2 is the area covered by Semi-natural systems and

C_3 is the area covered by High Naturalness systems.

3. Results

For the interpretation and the efficient, error-free usage of the results, some aspects have to be considered beforehand. In particular, these facets refer to:

- The meaning-interpretation of the NEI values: The NEI values range from 0 (totally artificial landscape) to 1 (landscapes totally covered by high-naturalness systems (e.g. forests)).
- Usual number of classes per NUTS-2 region and buffer's diameter: In the majority of the cases, NUTS-2 regions usually solely have 1 or 2 classes per buffer diameter. That means that the values of the Naturalness Evaluation Index can be more easily interpreted by the users of the outputs. In particular,

  NEI = 0 stands for totally artificial systems

  NEI = 0.3333 stands for totally agricultural systems

  NEI = 0.666 stands for totally semi-naturalness systems

  NEI = 1 stands for high naturalness systems

  0 < NEI < 0.3333 stands for both artificial and agricultural systems (if there are just two classes)

  0.3333 < NEI < 1 stands for both agricultural and high naturalness systems[1] (if there are just two classes)

- NUTS-2 regions with more number of classes than usual: Despite the fact that the majority of the NUTS-2 regions were represented by solely 1 or 2 classes, there are also NUTS-2 regions that their sample locations and their corresponding buffers were represented by 3 classes. These cases are the following:

  1. Latvia: 500 m diameter buffer – sample 16

[1] Semi-naturalness systems are not mentioned here as there was not such a class in the European NUTS-2 regions of the SOILGUARD Project.

1000 m diameter buffer – samples 8, 12, 16 - 19

2. Hungary: 1000 m diameter buffer – samples 16, 20, 27 and 28.

Naturalness Evaluation Index in EU NUTS-2 regions

All the results can be found in the provided excel files.

4. References

[1] CLC Product User Manual. (2018). https://land.copernicus.eu/. Retrieved June 30, 2023, from https://land.copernicus.eu/user-corner/technical-library/clc-product-user-manual

[2] Baiamonte, G., Domina, G., Raimondo, F. M., & Bazan, G. (2015). Agricultural landscapes and biodiversity conservation: a case study in Sicily (Italy). Biodiversity and Conservation, 24(13), 3201–3216. https://doi.org/10.1007/s10531-015-0950-4

Annex II. Small Woody Features 2015 dataset across the European NUTS-2 regions

1. Introducing the SWF 2015 dataset

SWFs are essential for a cornucopia of ecosystem services (ES), such as biodiversity and habitat connectivity. They are also fundamental actors in improving the quality of the air-soil-water system, managing greenhouse gas emissions, facilitating climate change adaption and carbon sequestration, regulating soil erosion and pollination as well as being vital for recreational, social and cultural reasons [1]. Last but not least, they are highly correlated with the green infrastructure strategy of the European Union (EU) and for monitoring the effectiveness of the Common Agricultural Policy (CAP).

The generation of the vector layer was based on the analysis of very-high resolution (VHR) data from the Copernicus Contribution Missions (CCMS) as seen in Figure 1 [1].

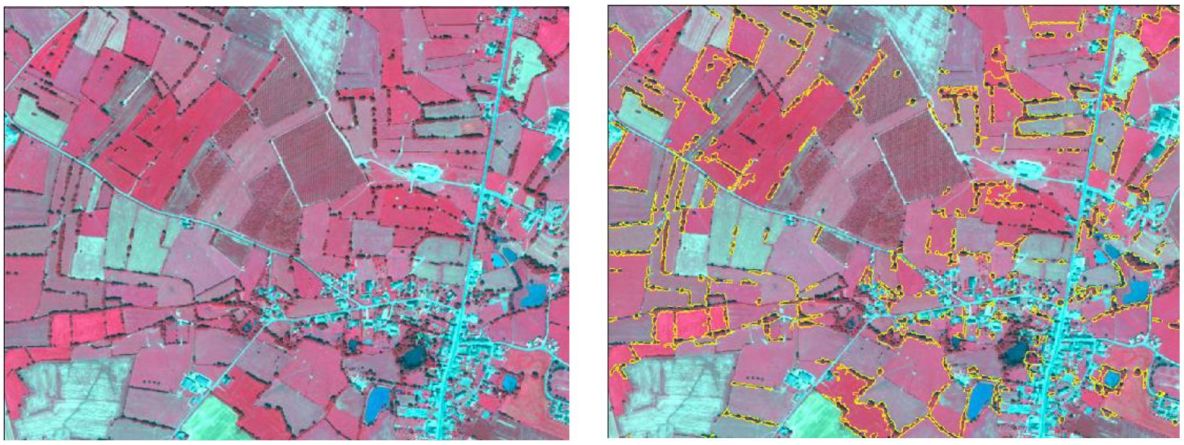


Figure 2. VHR image (2015) used as input data (left) and the SWF vector product (right with yellow outlines) [1].

While mapping the Small Woody Features, some elements were excluded due to the lack of reliable information and the limitations of the VHR data. As such, the manual of the dataset [1] differentiate the elements that were included and the ones that were excluded while generating the output dataset that refers to the SWF, as displayed in

Table 2.

Elements to be included in the SWF mapping (2015)	Elements to be excluded in the SWF mapping (2015)
Linear hedgerows and scrubs	Stone walls
Tree tows (e.g. along field boundaries)	Drainage ditches
Isolated / Scattered patches of trees	Grass margins
Additional Woody Features (neither linear nor patchy and with an area above 15000 m2)	Field boundaries without hedgerows or trees
	Any kind of 'grey' infrastructure such as roads
	Artificial tree rows like olive tree plantations, vineyards and orchards

Table 2. Elements of SWF mapping (2015) [1]

Elements to be included in the SWF mapping (2015)	Elements to be excluded in the SWF mapping (2015)[2]
Linear hedgerows and scrubs	Stone walls
Tree tows (e.g. along field boundaries)	Drainage ditches
Isolated / Scattered patches of trees	Grass margins
Additional Woody Features (neither linear nor patchy and with an area above 15000 m^2)	Field boundaries without hedgerows or trees
	Any kind of 'grey' infrastructure such as roads
	Artificial tree rows like olive tree plantations, vineyards and orchards[3]

For more information about the mapping rules that were applied to derive the homogeneous pan-European SWF dataset, the methodology of the layers generation etc., please follow the product specifications & user guidelines manual (https://land.copernicus.eu/user-corner/technical-library/hrl_lot5_d5-1_product-specification-document_i3-4_public-1.pdf) and the dataset's website https://land.copernicus.eu/pan-european/high-resolution-layers/small-woody-features/small-woody-features-2015.

2. Small Woody Features 2015 dataset

Materials & Methods

For acquiring information about the SWF 2015, the dataset was downloaded by the Copernicus data repository (Table 3).

Table 3. Materials used and relevant information.

Dataset	Coverage	Scale (m)	Temporal Reference	Source / Manual
Small Woody Features 2015	Pan - European	1:5000	2015	Copernicus/SWF_2015 / SWF_2015 Manual

The downloaded datasets were visually inspected, manipulated and processed into the ArcGIS Pro environment by using several geospatial tools such as project, merge, clip, intersect, buffer and summary statistics. Thenceforth, information about the SWF elements was extracted for the 7 European countries

[2] Despite the importance of the elements mentioned in this column, the type of Earth Observation (EO) data available to produce the layers did not allow to map this data reliably. Therefore, these elements were excluded from the product.

[3] It is noteworthy that a certain amount of commission is expected among the features that represent the artificial tree rows. These areas refer to the same land cover as natural woody features (equivalent spectral or textural information) but have a different soil use, which is difficult to extract in an automatic process.

and their corresponding NUTS-2 regions of the project. This information was statistically analyzed by using different metrics (see Results section).

3. Results

Statistical Analysis of the SWF 2015 dataset per NUTS-2 region across EU

Murcia, Spain

Total number of samples across the NUTS-2 region: **20**

Total area of the NUTS-2 region: **18253.1 km^2**

The following tables refer to Small Woody Features information in a local (sample sites) (Table 6 - Table 8), regional (NUTS-2 region) (Table 5) and national (whole Spain) (Table 4) level.

Table 4. Statistical analysis of the SWF dataset for the whole Spain.

Small Woody Features information for Spain						
Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	5339875	8401539328	1573.358801	0.114388	1182575.54	2278.091638
2	842138	491958973.5	584.178571	200.000113	4999.772796	631.710858
3	923080	11937027327	12931.7365	0.208489	49158501.04	102677.5605

Table 5. Statistical analysis of the SWF dataset for the Region de Murcia, Spain.

Small Woody Features information for Murcia, Spain						
Code	**Frequency**	**Total Area (m^2)**	**Mean Area (m^2)**	**Minimum Area (m^2)**	**Maximum Area (m^2)**	**STD Area (m^2)**
1	89856	169228634.1	1883.331488	0.236235	102101.5815	2840.31708
2	10845	7066547.873	651.595009	1.215235	4708.375119	655.286999
3	10013	77049929.97	7694.989511	4.93271	1716320.959	29037.05652

Table 6. Statistical analysis of the SWF dataset for the sample sites in Murcia, Spain. The SWF information was derived for a 200 m diameter buffer zone surrounding the sample sites.

200 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
6	1	1	1427.061965	1427.061965	1427.061965	1427.061965	0

The rest have no SWF in a 200 m buffer

Table 7. Statistical analysis of the SWF dataset for the sample sites in Murcia, Spain. The SWF information was derived for a 500 m diameter buffer zone surrounding the sample sites.

ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
3	1	6	5268.630204	878.105034	202.684573	1531.669961	655.791646
6	1	3	2488.083464	829.361155	426.944985	1427.061965	527.883095

500 m							
7				933.803171	933.803171	933.803171	0
22	1	2	10267.26904	5133.634521	3194.056142	7073.212899	2742.978048
22	3	1	619.348143	619.348143	619.348143	619.348143	0

The rest have no SWF in a 500 m buffer

Table 8. Statistical analysis of the SWF dataset for the sample sites in Murcia, Spain. The SWF information was derived for a 1000 m diameter buffer zone surrounding the sample sites.

1000 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1	7	13306.93913	1900.991304	43.743344	3953.499746	1237.056535
1	3	2	6733.555411	3366.777705	1595.492989	5138.062422	2504.974869
2	0	0	0	0	0	0	0
3	1	11	13164.77399	1196.797635	266.967904	2107.157038	667.258218
3	2	2	794.510434	397.255217	397.255217	397.255217	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	1	6	9308.059099	1551.343183	828.902382	2906.569058	817.320254
6	2	1	332.284274	332.284274	332.284274	332.284274	0
7	1	6	10254.25691	1709.042818	590.839937	3764.613248	1426.721761
7	2	2	952.919291	476.459645	9.214314	943.704976	660.784684
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	1	8	660.784684	517.857789	27.987702	826.131984	294.433449
10	3	2	6343.684113	3171.842056	3171.842056	3171.842056	0
22	1	9	55687.58024	6187.508916	40.358927	31201.24818	9608.988017
22	2	1	838.791161	838.791161	838.791161	838.791161	0
22	3	1	7538.486188	7538.486188	7538.486188	7538.486188	0
24	2	1	424.430751	424.430751	424.430751	424.430751	0

The rest have no SWF in a 1000 m buffer

Middle Jutland, Denmark

Total number of samples across the NUTS-2 region: **23**

Total area of the NUTS-2 region: **42567.6 km^2**

The following tables refer to Small Woody Features information in a local (sample sites) (Table 11 - Table 13), regional (NUTS-2 region) (Table 10) and national (whole Denmark) (Table 9) level.

Table 9. Statistical analysis of the SWF dataset for the whole Denmark.

Small Woody Features information for Denmark						
Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area ()	STD Area (m^2)

1	936487	5508939454	5882.558384	299.767232	1652620.484	10653.92932
2	85450	280001728.3	3276.790267	597.208526	17321.79368	2829.882816
3	144141	4613215655	32004.88171	4480.784473	5082430.174	65002.67721

Table 10. Statistical analysis of the SWF dataset for Middle Jutland, Denmark.

Small Woody Features information for Middle Jutland, Denmark						
Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	287004	1678622174	5848.776233	0.172734	1618060.487	9972.362095
2	23908	86734590.59	3627.848025	0.452991	16239.0236	2952.170708
3	49192	1875350890	38123.08688	0.448771	5082430.174	82551.48745

Table 11. Statistical analysis of the SWF dataset for the sample sites in Middle Jutland, Denmark. The SWF information was derived for a 200 m diameter buffer zone surrounding the sample sites.

200 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1	2	5162.417735	2581.208867	2581.208867	5137.444387	3615.062941
2	1	2	1429.594747	714.797374	649.796642	779.798105	91.924916
2	3	1	10470.20329	10470.20329	10470.20329	10470.20329	0
3	1	1	2242.049395	2242.049395	2242.049395	2242.049395	0
9	1	2	5476.397852	2738.198926	355.475905	5120.921947	3369.679211
12	1	1	1238.094531	1238.094531	1238.094531	1238.094531	0
15	1	2	4765.361896	2382.680948	435.628173	4329.733722	2753.54844
15	3	1	4430.013321	4430.013321	4430.013321	4430.013321	0
16	1	1	2153.22354	2153.22354	2153.22354	2153.22354	0
18	1	2	2968.292709	1484.146355	253.360983	2714.931726	1740.593364
18	3	1	1790.503986	1790.503986	1790.503986	1790.503986	0
22	1	3	2091.322218	697.107406	1.802295	1293.327643	651.439015
22	3	1	20618.00898	20618.00898	20618.00898	20618.00898	0
23	1	2	2173.185836	1086.592918	295.049293	1878.136543	1119.41173
24	1	1	295.049293	295.049293	295.049293	295.049293	0
25	1	1	1736.694685	1736.694685	1736.694685	1736.694685	0
26	1	1	114.813257	114.813257	114.813257	114.813257	0
27	1	1	1341.360877	1341.360877	1341.360877	1341.360877	0

The rest have no SWF in a 200 m buffer

Table 12. Statistical analysis of the SWF dataset for the sample sites in Middle Jutland, Denmark. The SWF information was derived for a 500 m diameter buffer zone surrounding the sample sites.

500 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)

1	1	2	15834.68596	7917.342978	2370.719285	13463.96667	7844.110453
2	1	2	3959.188388	1979.594194	1194.260075	2764.928313	1110.630163
2	3	1	52920.81098	52920.81098	52920.81098	52920.81098	0
3	1	4	26719.52925	6679.882312	923.137462	18733.71878	8230.003605
3	3	1	1996.778754	1996.778754	1996.778754	1996.778754	0
9	1	2	14763.75423	7381.877117	6460.986026	8302.768208	1302.33667
11	1	1	190.809545	190.809545	190.809545	190.809545	0
12	1	1	2661.77781	2661.77781	2661.77781	2661.77781	0
15	1	4	9611.858381	2402.964595	824.269926	4329.733722	1539.289758
15	3	1	17037.47342	17037.47342	17037.47342	17037.47342	0
16	1	2	4341.247293	2170.623646	1608.049494	2733.197799	795.599996
17	1	2	1408.222185	704.111092	114.077879	1294.144306	834.432973
18	1	6	16343.52564	2723.92094	301.766704	6744.862097	2366.33761
18	3	1	38032.56063	38032.56063	38032.56063	38032.56063	0
19	1	2	3296.631316	1648.315658	235.727476	3060.903839	1997.701364
21	1	1	73.860448	73.860448	73.860448	73.860448	0
22	1	5	12234.56176	2446.912352	1293.327643	4269.047058	1107.141902
22	2	1	3455.797104	3455.797104	3455.797104	3455.797104	0
22	3	1	97288.34715	97288.34715	97288.34715	97288.34715	0
23	1	5	5605.253265	1121.050653	258.796066	2850.028647	1085.435914
24	1	2	3558.702399	1779.3512	708.673752	2850.028647	1514.166567
25	1	2	6577.281111	3288.640555	1736.694685	4840.586425	2194.782897
26	1	3	8398.090827	2799.363609	1.778743	6866.173976	3603.911834
26	3	1	4008.474749	4008.474749	4008.474749	4008.474749	0
27	1	4	8673.450108	2168.362527	840.604342	4982.238747	1904.159997
27	2	2	5587.758626	2793.879313	652.533995	4935.224631	3028.319591
28	1	1	374.416206	374.416206	374.416206	374.416206	0

The rest have no SWF in a 500 m buffer

Table 13. Statistical analysis of the SWF dataset for the sample sites in Middle Jutland, Denmark. The SWF information was derived for a 1000 m diameter buffer zone surrounding the sample sites.

1000 m

ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1	3	17580.35489	5860.118295	19.91549688	15189.7201	8164.721186
2	1	2	3959.188388	1979.594194	1194.260075	2764.928313	1110.630163
2	3	1	135489.0628	135489.0628	135489.0628	135489.0628	0
3	1	12	50219.27786	4184.939821	301.4222353	18733.72243	5276.266105
3	2	1	1298.972683	1298.972683	1298.972683	1298.972683	0
3	3	3	96528.59228	32176.19743	20792.10704	39764.86147	10039.68617
9	1	4	26518.7697	6629.692424	3579.634808	8302.768208	2199.973333
10	1	2	4812.40534	2406.20267	1370.914927	3441.490414	1464.117968
11	1	3	5585.48288	1861.827627	889.7469675	3412.905933	1357.575595
12	1	5	12506.25004	2501.250008	114.4643483	4926.455351	2069.335441

15	1	9	63793.47166	7088.163518	1252.966138	43351.29511	13631.44782
15	2	1	834.1288497	834.1288497	834.1288497	834.1288497	0
15	3	1	17037.47342	17037.47342	17037.47342	17037.47342	0
16	1	4	11937.50685	2984.376712	1608.049494	5492.1798	1734.105583
17	1	3	7949.424061	2649.80802	1294.144306	3817.505121	1272.138372
17	3	1	194.0294699	194.0294699	194.0294699	194.0294699	0
18	1	7	19698.17895	2814.025564	736.9331923	6780.565788	2070.585413
18	2	1	824.4845176	824.4845176	824.4845176	824.4845176	0
18	3	1	41242.50331	41242.50331	41242.50331	41242.50331	0
19	1	6	18355.88502	3059.314169	22.64982417	4981.696295	1871.253292
19	2	2	1456.165071	728.0825355	702.4586821	753.7063888	36.23760091
21	1	6	14779.30472	2463.217453	508.9054214	8158.807313	2865.917622
21	2	1	773.4264139	773.4264139	773.4264139	773.4264139	0
22	1	11	28764.72569	2614.975063	118.0649452	5021.605442	1599.931936
22	2	1	3455.797104	3455.797104	3455.797104	3455.797104	0
22	3	2	199525.8789	99762.93945	6682.844976	192843.0339	131635.132
23	1	10	26972.19978	2697.219978	379.4867886	6866.941547	1971.343023
23	3	3	9662.455731	3220.818577	252.9122824	6652.936824	3225.166366
24	1	11	29200.72434	2654.611304	379.4867886	6866.941547	1866.509533
24	3	1	6652.936824	6652.936824	6652.936824	6652.936824	0
25	1	11	16813.09219	3362.618438	1590.279302	5666.486493	1998.959681
25	2	1	1511.206837	1511.206837	1511.206837	1511.206837	0
25	3	2	5094.448786	2547.224393	18.21471491	5076.234071	3576.559786
26	1	5	14009.81654	2801.963307	49.65646845	6866.173976	2688.219134
26	3	1	9662.455731	3220.818577	252.9122824	6652.936824	3225.166366
27	1	11	27915.1269	2537.738809	512.8318676	8149.806542	2355.663545
27	2	5	7680.866335	1536.173267	652.5339946	2681.144322	891.3901368
28	1	15	49970.85297	3331.390198	73.31590983	17184.51011	4427.75974
28	2	3	4166.125284	1388.708428	652.5339946	2681.144322	1122.89139
28	3	1	13760.57154	13760.57154	13760.57154	13760.57154	0

The rest have no SWF in a 1000 m buffer

South Denmark

Total number of samples across the NUTS-2 region: **7**

Total area of the NUTS-2 region: **37533.5 km²**

The following tables refer to Small Woody Features information in a local (sample sites) (Table 15 - Table 17), regional (NUTS-2 region) (Table 14) and national (whole Denmark) (Table 9) level.

Table 14. Statistical analysis of the SWF dataset for the South Denmark

Small Woody Features information for South Denmark						
Code	Frequency	Total Area (m²)	Mean Area (m²)	Minimum Area (m²)	Maximum Area (m²)	STD Area (m²)
1	253003	1506833381	5955.792545	0.003069	1228579.512	11979.96595
2	23274	73803956.86	3171.090352	36.808692	15566.59064	2724.053367
3	39291	1107171996	28178.76856	2.164198	1813140.197	47900.8739

Table 15. Statistical analysis of the SWF dataset for the sample sites in South Denmark. The SWF information was derived for a 200 m diameter buffer zone surrounding the sample sites.

200 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
5	1	2	2228.410561	1114.20528	443.064041	1785.34652	949.137043
13	1	1	2874.251924	2874.251924	2874.251924	2874.251924	0
14	1	2	4091.9103	2045.95515	765.389251	3326.521049	1810.993662

The rest have no SWF in a 200 m buffer

Table 16. Statistical analysis of the SWF dataset for the sample sites in South Denmark. The SWF information was derived for a 500 m diameter buffer zone surrounding the sample sites.

500 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
5	1	4	6056.776186	1514.194047	1272.615588	1785.34652	210.941775
13	1	6	10041.83106	1673.638509	208.633389	6468.64178	2366.414939
14	1	2	14357.89507	7178.947533	2900.009109	11457.88596	6051.332752

The rest have no SWF in a 500 m buffer

Table 17. Statistical analysis of the SWF dataset for the sample sites in South Denmark. The SWF information was derived for a 1000 m diameter buffer zone surrounding the sample sites.

1000 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
5	1	11	38485.43171	3498.67561	928.042986	12222.99468	3740.922883
5	3	3	29225.79726	9741.932418	1782.693464	16709.65748	7512.715124
6	1	1	2193.132205	2193.132205	2193.132205	2193.132205	0
13	1	10	55284.04707	5528.404707	787.575254	12541.34093	3652.365283
13	2	1	3841.969382	3841.969382	3841.969382	3841.969382	0
14	1	8	41998.27004	5249.783755	512.744971	12252.53981	4847.266728
14	2	1	976.042487	976.042487	976.042487	976.042487	0
14	3	2	37422.57275	18711.28638	5698.632366	31723.94039	18402.67178

The rest have no SWF in a 1000 m buffer

Southern Ireland

Total number of samples across the NUTS-2 region: **30**

Total area of the NUTS-2 region: **79388.5 km^2**

The following tables refer to Small Woody Features information in a local (sample sites) (Table 20 - Table 22), regional (NUTS-2 region) (Table 19) and national (whole Ireland) (Table 18) level.

Table 18. Statistical analysis of the SWF dataset for the whole Ireland.

Small Woody Features information for Ireland						
Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1724799	13866005418	8039.200752	275.709344	7274222.68	24265.50142
2	178011	492288733.8	2765.496142	516.277897	15145.96035	2277.643455
3	270881	9182299167	33897.90781	3864.753516	7093979.319	90865.33102

Table 19. Statistical analysis of the SWF dataset for Southern Ireland.

Small Woody Features information for Southern Ireland						
Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	741928	5977099617	8056.172051	0.000409	7274222.68	25636.51466
2	77035	209498492.6	2719.523497	0.004589	13826.47108	2199.161056
3	119080	3746292964	31460.30369	1.396185	6563234.991	81473.36351

Table 20. Statistical analysis of the SWF dataset for the sample sites in Southern Ireland. The SWF information was derived for a 200 m diameter buffer zone surrounding the sample sites.

		200 m					
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1	1	552.361648	552.361648	552.361648	552.361648	0
2	1	1	4178.652629	4178.652629	4178.652629	4178.652629	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	1	2	4685.444272	2342.722136	1846.527973	2838.916299	701.724515
6	1	2	3093.323463	1546.661731	254.407164	2838.916299	1827.523935
7	1	1	194.174221	194.174221	194.174221	194.174221	0
8	1	2	2737.169001	1368.5845	1058.522912	1678.646089	438.493304
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	1	2	1994.304575	997.152287	872.387359	1121.917216	176.444254
13	3	1	323.071241	323.071241	323.071241	323.071241	0
14	1	2	4103.431719	2051.715859	872.387359	3231.04436	1667.82236
14	3	2	5214.758695	2607.379347	323.071241	4891.687454	3230.499504
15	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0
17	1	3	3224.584069	1074.861356	0.151783	2931.883932	1614.860233
18	1	3	6706.068656	2235.356219	0.151783	3774.032941	1981.011835

19	1	1	409.663396	409.663396	409.663396	409.663396	0
20	1	2	800.680852	400.340426	391.017456	409.663396	13.184671
21	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0
23	1	1	1666.478224	1666.478224	1666.478224	1666.478224	0
23	3	1	13494.59665	13494.59665	13494.59665	13494.59665	0
24	1	1	1316.354955	1316.354955	1316.354955	1316.354955	0
24	3	1	11439.58495	11439.58495	11439.58495	11439.58495	0
25	3	1	10377.3983	10377.3983	10377.3983	10377.3983	0
26	1	1	572.733578	572.733578	572.733578	572.733578	0
26	3	2	15181.79703	7590.898514	1822.408212	13359.38882	8157.87722
27	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0
29	1	2	549.727275	274.863638	216.627415	333.099861	82.358457
30	1	1	101.718683	101.718683	101.718683	101.718683	0

The rest have no SWF in a 200 m buffer

Table 21. Statistical analysis of the SWF dataset for the sample sites in Southern Ireland. The SWF information was derived for a 500 m diameter buffer zone surrounding the sample sites.

500 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1	5	4555.068241	911.0136483	223.7993017	2821.139087	1087.046481
2	1	5	16929.27675	3385.855349	223.7993017	14349.38782	6151.298207
3	1	2	1822.700108	911.3500539	735.6294969	1087.070611	248.5063949
4	1	2	1033.642949	516.8214745	298.0134522	735.6294969	309.4412728
5	1	7	21782.43527	3111.776467	11.58293721	13918.53478	4977.951836
5	2	1	3512.014785	3512.014785	3512.014785	3512.014785	0
5	3	3	9131.523174	3043.841058	18.56091879	5013.443016	2659.521406
6	1	7	23008.69623	3286.956604	62.49420741	13918.53478	5055.988524
6	2	1	3512.014785	3512.014785	3512.014785	3512.014785	0
6	3	1	4099.519239	4099.519239	4099.519239	4099.519239	0
7	1	4	8284.182715	2071.045679	148.6611875	4740.299084	2001.634594
8	1	7	11159.4778	1594.211114	455.4303011	4929.985727	1535.152027
8	2	1	789.6445797	789.6445797	789.6445797	789.6445797	0
9	0	0	0	0	0	0	0
10	1	1	516.8149946	516.8149946	516.8149946	516.8149946	0
11	1	1	469.1994886	469.1994886	469.1994886	469.1994886	0
12	1	3	1284.097477	428.0324922	156.9825737	657.9154144	252.9910451
13	1	6	10741.11783	1790.186306	161.2434815	7576.926742	2880.29383
13	3	3	8164.074645	2721.358215	806.8589497	5428.081357	2410.319094
14	1	7	15715.01953	2245.00279	161.2434815	7576.926742	2572.426394
14	3	4	10915.38363	2728.845906	22.03999881	5428.081357	2709.852143
15	1	2	613.5254917	306.7627458	125.1445283	488.3809633	256.8469464

15	3	2	3905.349149	1952.674574	968.9171073	2936.432042	1391.243152
16	1	1	488.3809633	488.3809633	488.3809633	488.3809633	0
16	3	1	968.9171073	968.9171073	968.9171073	968.9171073	0
17	1	6	30922.67156	5153.778593	377.0388155	18039.94233	6694.684143
18	1	6	31867.09424	5311.182373	712.5390336	18039.94233	6565.916323
19	1	3	13633.16623	4544.388743	100.2652607	11730.62991	6281.381421
19	3	2	390.2509523	195.1254761	0.381647571	389.8693047	275.4093635
20	1	4	13520.32581	3380.081453	100.2652607	11730.62991	5606.807428
20	3	1	0.381647571	0.381647571	0.381647571	0.381647571	0
21	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0
23	1	7	9935.45717	1419.351024	1.902501298	2675.247848	1070.699207
23	2	1	3.707887857	3.707887857	3.707887857	3.707887857	0
23	3	3	48770.06582	16256.68861	3233.561867	35167.8545	16761.60838
24	1	2	4254.464022	2127.232011	1150.346505	3104.117517	1381.524731
24	3	1	67778.91843	67778.91843	67778.91843	67778.91843	0
25	1	2	3161.670838	1580.835419	1049.045654	2112.625184	752.0642981
25	3	3	82166.65688	27388.88563	12014.753	35615.892	13325.33304
26	1	4	4125.240634	1031.310159	172.788367	2278.624188	905.2497447
26	3	4	74091.11035	18522.77759	3615.219885	35615.892	13849.76473
27	1	3	3062.324307	1020.774769	722.0291989	1427.379443	364.8354857
27	3	3	22746.95654	7582.318846	1673.781488	16507.34357	7863.398629
28	1	5	4658.78855	931.7577101	78.67067071	2209.994181	781.7656773
28	3	1	4565.831484	4565.831484	4565.831484	4565.831484	0
29	1	5	20743.47779	4148.695558	200.3861026	6431.084932	2607.122253
30	1	6	15303.49778	2550.582964	23.40622104	5632.741161	2142.435004
30	2	1	997.0386037	997.0386037	997.0386037	997.0386037	0
30	3	4	16438.39614	4109.599036	28.8786911	6163.292909	2801.154437

The rest have no SWF in a 500 m buffer

Table 22. Statistical analysis of the SWF dataset for the sample sites in Southern Ireland. The SWF information was derived for a 1000 m diameter buffer zone surrounding the sample sites.

1000 m							
ID	Code	Frequency	Total Area (m^2)	Mean Area (m^2)	Minimum Area (m^2)	Maximum Area (m^2)	STD Area (m^2)
1	1	11	46743.2547	4249.386791	34.66831244	16407.10432	5889.415127
2	1	8	38537.38257	4817.172821	577.9122495	16407.10432	6665.406549
2	2	1	588.5305879	588.5305879	588.5305879	588.5305879	0
3	1	5	7840.258728	1568.051746	1106.983889	1962.143468	329.1684708
4	1	6	8788.163269	1464.693878	947.904541	1962.143468	388.3021997
5	1	21	105117.7107	5005.60527	12.18834234	38733.33279	8755.607186
5	2	3	10606.21035	3535.403451	1893.604578	5200.59099	1653.617264
5	3	8	108614.1558	13576.76948	1170.097181	32130.09904	10378.51968
6	1	19	99953.3976	5260.705137	12.18834234	38733.33279	9073.478281

6	2	3	9331.973387	3110.657796	619.3676113	5200.59099	2316.833504
6	3	7	95686.68345	13669.52621	157.4850124	32130.09904	11561.20533
7	1	24	68341.92914	2847.580381	2.925997177	10284.16105	2823.535656
7	2	1	3962.110584	3962.110584	3962.110584	3962.110584	0
7	3	1	2031.039397	2031.039397	2031.039397	2031.039397	0
8	1	17	68921.01136	4054.177139	94.29642391	16109.47371	4165.167785
8	2	1	3962.110584	3962.110584	3962.110584	3962.110584	0
9	1	4	6858.210694	1714.552673	351.7494262	3190.102117	1186.502347
9	2	1	2605.913464	2605.913464	2605.913464	2605.913464	0
9	3	1	10435.82193	10435.82193	10435.82193	10435.82193	0
10	1	6	10251.83108	1708.638514	47.33472964	7457.70526	2850.687834
10	3	2	14537.14977	7268.574885	4633.89268	9903.25709	3726.003306
11	1	8	30246.79015	3780.848769	477.0691008	13129.79595	4320.867115
12	1	8	32566.30839	4070.788549	244.7097619	13129.79595	4222.170059
13	1	16	69296.57842	4331.036151	842.5124212	11988.11609	3466.726788
13	3	4	59267.5154	14816.87885	5428.081357	25182.75953	8084.992344
14	1	17	68810.46776	4047.674574	289.7835675	11988.11609	3542.464986
14	3	6	66635.84298	11105.97383	1671.540095	25182.75953	8595.957245
15	1	5	10394.43017	2078.886035	202.3797022	4123.185508	1482.872714
15	3	1	14240.10082	14240.10082	14240.10082	14240.10082	0
16	1	4	10192.05047	2548.012618	1274.130386	4123.185508	1210.241151
16	3	1	14240.10082	14240.10082	14240.10082	14240.10082	0
17	1	20	110665.1877	5533.259385	126.1054788	35209.7535	9028.716102
17	3	2	712.5599184	356.2799592	100.5531465	612.006772	361.6523268
18	1	19	113366.0959	5966.636626	126.1054788	35209.7535	9130.986308
18	3	2	806.7915925	403.3957962	194.7848205	612.006772	295.0204711
19	1	14	66638.19579	4759.871128	552.2594417	16297.40126	5055.085326
19	3	3	25187.22116	8395.740386	626.6444456	17529.91586	8533.916319
20	1	14	66446.53086	4746.180775	279.0382591	16297.40126	5069.103069
20	3	4	26947.8305	6736.957625	626.6444456	17529.91586	7717.38683
21	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0
23	1	18	55463.14089	3081.285605	0.112997447	15605.61455	3963.635022
23	2	1	4846.365484	4846.365484	4846.365484	4846.365484	0
23	3	7	92119.71897	13159.95985	348.0004894	35424.64285	13581.83582
24	1	7	36330.99187	5190.141696	1150.346505	10798.90495	3535.342986
24	3	4	266097.9785	66524.49462	348.0004894	256916.9611	126942.8037
25	1	26	68661.10041	2640.811554	56.02877022	14904.95322	3204.793917
25	2	1	754.51283	754.51283	754.51283	754.51283	0
25	3	8	239589.5668	29948.69585	149.8547127	128974.9362	45050.53217
26	1	19	59006.92527	3105.627646	7.648221712	17697.2142	4059.089128
26	2	1	754.51283	754.51283	754.51283	754.51283	0
26	3	6	215885.7724	35980.96207	1398.638054	128974.9362	51340.46086
27	1	16	57275.32428	3579.707768	117.571672	10684.47417	3188.814234
27	2	2	2806.741452	1403.370726	59.22412758	2747.517325	1900.91035

27	3	4	121303.3951	30325.84876	10010.36735	46437.97208	15377.48746
28	1	17	52468.90714	3086.406302	117.2112995	10684.47417	2956.643757
28	2	1	59.22412758	59.22412758	59.22412758	59.22412758	0
28	3	2	82623.68756	41311.84378	36185.71548	46437.97208	7249.440166
29	1	27	103418.3642	3830.309787	195.9328435	12535.55286	3033.709492
29	2	1	2752.550742	2752.550742	2752.550742	2752.550742	0
29	3	4	22682.81057	5670.702642	180.9770321	9955.859635	4828.333363
30	1	23	67816.65371	2948.550161	2.416658734	10668.15807	2712.298075
30	2	1	2105.542305	2105.542305	2105.542305	2105.542305	0
30	3	7	122578.3294	17511.18991	5677.658631	76052.45456	25880.4006

The rest have no SWF in a 1000 m buffer

West Flanders, Belgium

Total number of samples across the NUTS-2 region: **20**

Total area of the NUTS-2 region: **7995.82 km²**

The results of the statistical analysis can be found in the excel file.

Latvia

Total number of samples across the NUTS-2 region: **20**

Total area of the NUTS-2 region: **215510 km²**

The results of the statistical analysis can be found in the excel file.

South Transdanubia, Hungary

Total number of samples across the NUTS-2 region: **20**

Total area of the NUTS-2 region: **29774.8 km²**

The results of the statistical analysis can be found in the excel file.

Western Finland

Total number of samples across the NUTS-2 region: **24**

Total area of the NUTS-2 region: **298701 km²**

The results of the statistical analysis can be found in the excel file.

4. References

[1] Copernicus Land Monitoring Service – High Resolution Layer Small Woody Features – 2015 reference year Product Specifications & User Guidelines. (2019). https://land.copernicus.eu/. Retrieved May 15, 2023, from https://land.copernicus.eu/user-corner/technical-library/hrl_lot5_d5-1_product-specification-document_i3-4_public-1.pdf

Annex III. Extended methods and results of molecular analysis of biodiversity

The biodiversity of each sample belonging to the cross-biome network of sites was assessed through DNA metabarcoding with taxonomy markers for prokaryotes, fungi and rest of eukaryotes. Metabarcoding is widely used methodological approach based on amplification and sequencing that can allow us to tackle the most relevant kingdoms of the soil community (Geisen et al., 2018). However, the selection of primers for the amplification will strongly affect the subsequent biodiversity results (He et al., 2023; Li et al., 2020). In this project, we targeted three gene regions indicated by the SOILBON European soil initiative (https://www.globalsoilbiodiversity.org/soilbon): 16S V3-V4 region for prokaryotes, ITS2 region for fungi and 18S V4 region with universal primers for Eukaryotes (Table S1). The latter allows to obtain data for protists, nematodes, arthropods and earthworms. In addition to that, we also completed the sequencing of other regions, which were described as more specific for taxa belonging to protists (18S V4-V5 region), nematodes (18S V6-V8 region), arthropods (COI region) and earthworms (mit 16S region, Table 2). In this report we will focus on the datasets generated from primers determined by SOILBON, namely for the 16S, ITS2 and the universal 18S region (see section 3.2).

Table S1. *Primers selected to target specific taxon groups.*

Target group	Primers (sequence, 5'-3') / Region	Region	Ref.
Bacteria/archeae	341F (CCTAYGGGDBGCWSCAG)	V3-V4 16S	(Frey et al., 2016)
	806R (GGACTACNVGGGTHTCTAAT)		
Fungi	ITS3ngs (CANCGATGAAGAACGYRG)	ITS2	(Tedersoo and Lindahl, 2016)
	ITS4ngs (CCTSCSCTTANTDATATGC)		
Eukaryotes	Euk575Fngs (ASCYGYGGTAAYWCCAGC)	V4 18S	(Guerra et al., 2021)
	Euk895Rngs (TCHNHGNATTTCACCNCT)		
Earthworms	ewD (ATTCGGTTGGGGCGACC)	mit 16S	(Bienert et al., 2012)
	ewE (CTGTTATCCCTAAGGTAGCTT)		
Protist	616*f (TTAAARVGYTCGTAGTYG)	V4-V5 18S	(Hugerth et al., 2014)
	1132r (CCGTCAATTHCTTYAART)		
Nematodes	Nemf (GGGGAAGTATGGTTGCAAA)	V6-V8 18S	(Sikder et al., 2020)
	18Sr2b (TACAAAGGGCAGGGACGTAAT)		
Arthropods	BF3 (CCHGAYATRGCHTTYCCHCG)	COI	(Elbrecht et al., 2019)
	BR2 (TCDGGRTGNCCRAARAAYCA)		

When received the DNA, amplicon libraries were generated targeting the V3-V4 regions (341F/806R) of the 16S ribosomal RNA gene for Prokaryotes, the region 2 of internal transcriber spacer (ITS3ngs/ITS4ngs) region for Fungi and V4 region of the 18S ribosomal gene (Euk575Fngs/Euk985Rngs) for the rest of Eukaryotes. The sequencing was performed on the Illumina NextSeq platform PE300 60M reads, following the recommendations of Illumina Inc. Sequencing in Genome Québec Inc. (Centre d'expertise et de services Génome Québec, Montréal, Québec, Canada). All PCR conditions during library preparation were based on previous literature description, performing 38 cycles for prokaryotes, 33 for fungi, 35 for eukaryotes.

All samples collected were amplified and sequenced successfully for the three target markers used, with a very similar global output of high-quality reads (~30M reads per amplicon, Supplementary Table S2). This raw sequencing data was processed based on the same bioinformatic pipeline that WP3 to give robustness across WPs (see Deliverable 3.2), largely based on VSEARCH (Rognes et al., 2016) as previously described (Longepierre et al., 2021). Briefly, the trimming of primers from paired-end reads was performed

with Cutadapt; the merging of paired-end reads was done with VSEARCH; quality filtering by maximum expected error of 1 was carried out with VSEARCH and delineation n of sequences into amplicon sequence variants (ASVs) with minsize of 8; removal of chimeras (VSEARCH, Rognes et al. 2016); target verification using Metaxa2, for the 16S and 18S rRNA genes and ITSx (Bengtsson-Palme et al., 2015) for the ITS2 sequences using VSEARCH with settings, maxaccepts 100, maxhits 1, and a minimum identity of 97%. Taxonomic classification of each verified ASV sequence was performed by running the SINTAX algorithm implemented in VSEARCH against the SILVA v.138 database **(Pruesse et al., 2007)** for the 16S rRNA gene sequences (bacteria and archaea), against the UNITE v.8.3 database (**Abarenkov et al., 2010**) for the ITS2 sequences (fungi), against the PR2 v5.0 database for 18S rRNA gene sequences, against MIDORI GB257 for the 16S mitochondrial sequences (annelids) and against BOLD Feb-2023 for the COI sequences (arthropods) using a bootstrap cutoff of 0.8. For the three datasets, we normalized the reads numbers of all samples and further data analyses under R software, as is explained above.

Subtask 2.2.1

Virome, antibiotic resistance genes (ARGs) and functional analysis

233 samples were sent for the analysis in February, currently under sequencing with an expected 100M reads on average per sample. In the meantime, optimization of the virome pipeline or antibiotic resistance genes identification is under process in the framework of another project.

Soil arthropods primers investigation (COI gene):

Several in silico tests have been performed, and three pairs perform better than the rest:

Table S2. *Primer pairs identified for arthropods metabarcoding study. Only 1 has previously been tested in soil samples.*

	SEQUENCE (5'- 3')	**size**	**Nº MATCH: In silico all BOLD**	**Nº MATCH: In silico NCBI+BOLD databases (only arth)**	**REFERENCE**	**Target**
fwhF2	GGDACWGGWTGAACWGTWTAYCCHCC	142 bp	494,236 (ALL)	287,345 (ARTH)	Leese et. al 2020	COI GENE
EPTDr2n	CAAACAAATARDGGTATTCGDTY		492,018 (ARTH)			
IllBF	CCN GAY ATR GCN TTY CCN CG	315 bp	31,101 (ALL)	1,123 (ARTH)	Lentendu et al. 2022	
ArR5	GTR ATN GCN CCN GCN ARN AC		1,371 (ARTH)			
B	CCIGAYATRGCITTYCCICG	315 bp	31,101 (ALL)	1,123 (ARTH)	Teresita M. Porter at al. 2019	
E	GTRATIGCICCIGCIARIAC		1,371 (ARTH)			

However, after discuss with the sequencing company we decided to try the following: mICOIintF - GGWACWGGWTGAACWGTWTAYCCYCC and jgHCO2198 – TAIACYTCIGGRTGICCRAARAAYCA; which

general have successful detection for arthropods in complex samples (Leray et al. 2013, Isabwe et al. 2022). Data hava been received during July and are currently being processed.

Primers comparison for metabarcoding biodiversity assessment of eukaryotes

Global eukaryotes

18SV4 region targeted encompass the higher diversity values in most of the regions and a global higher number of genera identified. We can confirm that these primers are the best to global assesment of eukaryotes biodiversity in different types of agroecosystems.

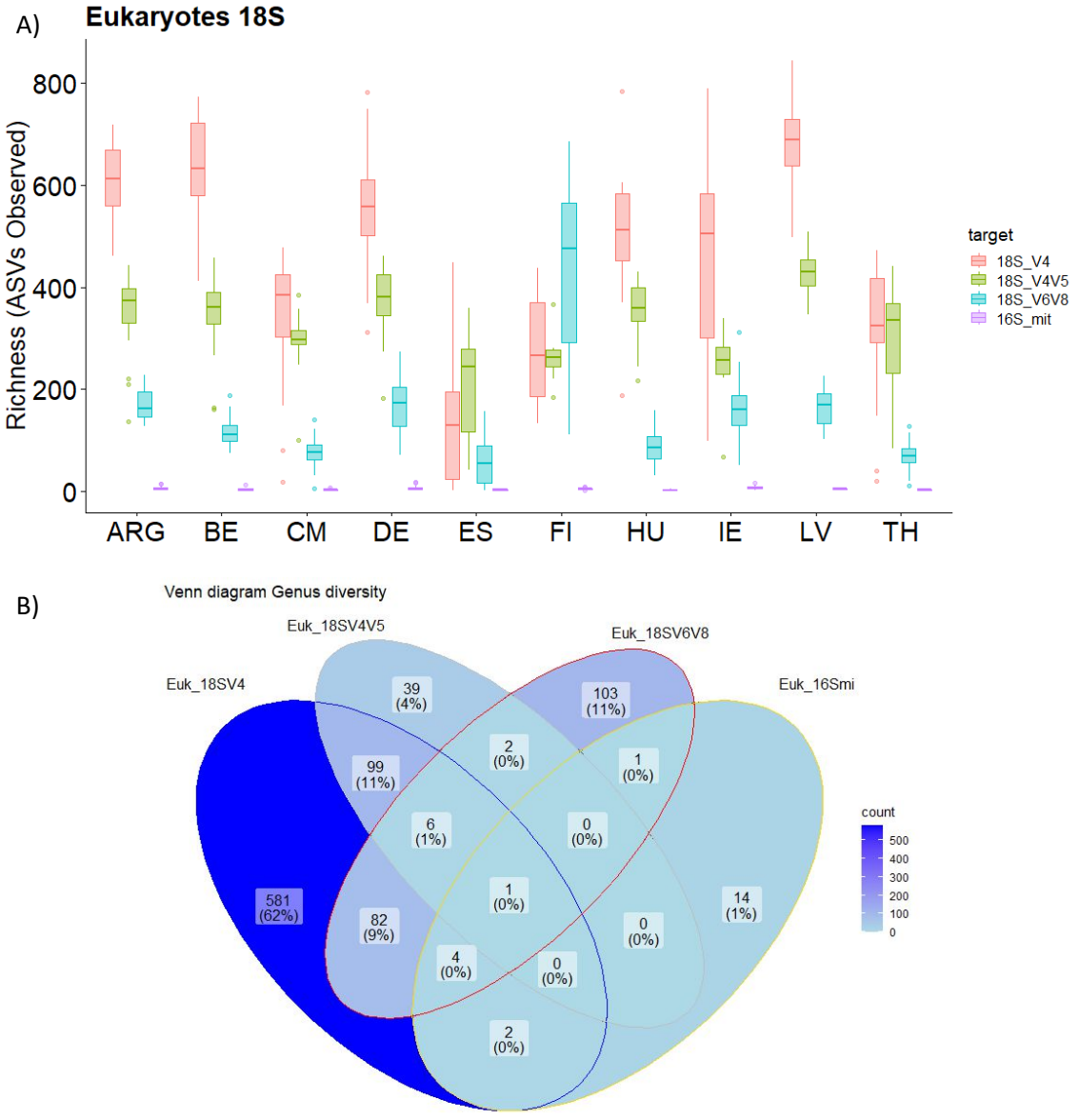


Figure S1. A) Alfa diversity distribution analysis for the different target regions/genes related to eukaryotes, divided by country. For the combination of country x target, the observed ASVs are represented by boxplots. B) Total of shared and non-shared genera between sequencing datasets.

Protist

Higher ASVs number observed in 18SV4 for all regions than in 18SV4V5, also in terms of genus. Beta-diversity from both datasets allows us to observe interesting patterns between regions (not shown in this document), but we must consider that 30 more samples are lost in the V4V5 dataset because of non-enough protist reads identified. All in all, we suggest using the 18SV4 gene for this target group based on the comparisons performed under this project.

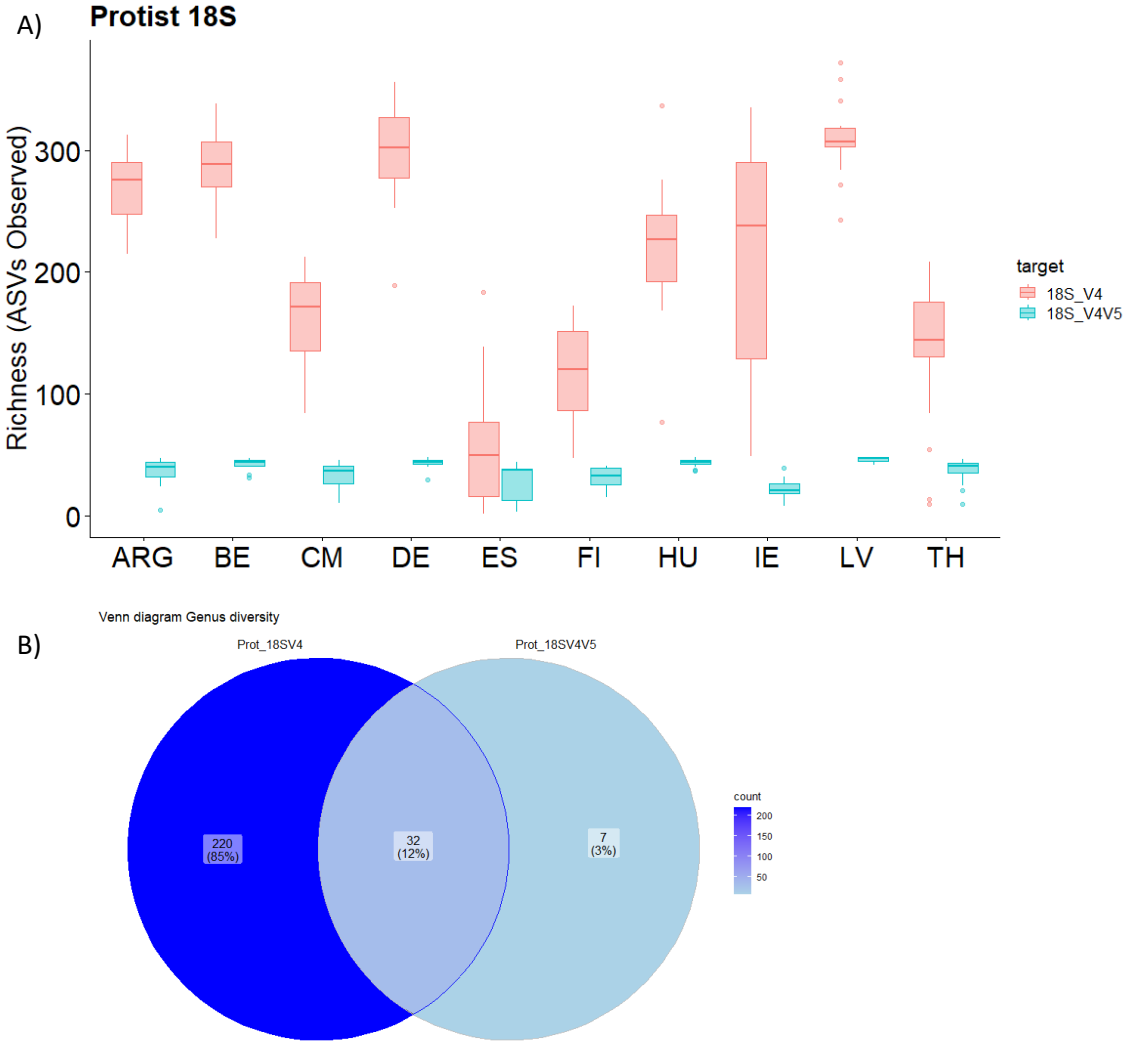


Figure S2. A) Alfa diversity distribution analysis for the different target regions/genes related to protists, divided by country. For the combination of country x target, the observed ASVs are represented by boxplots. B) Total of shared and non-shared genera between sequencing datasets.

Nematodes

Higher ASVs number observed in 18SV6V8 for all regions than in 18SV4, also in terms of genus. Moreover, higher number of reads per ASV was captured in the 18SV6V8 sequencing dataset, which allow us to select 1,200 depths as minimun read count instead of 50. Thus, we select the 18SV6V8 dataset to further analysis of Nematodes.

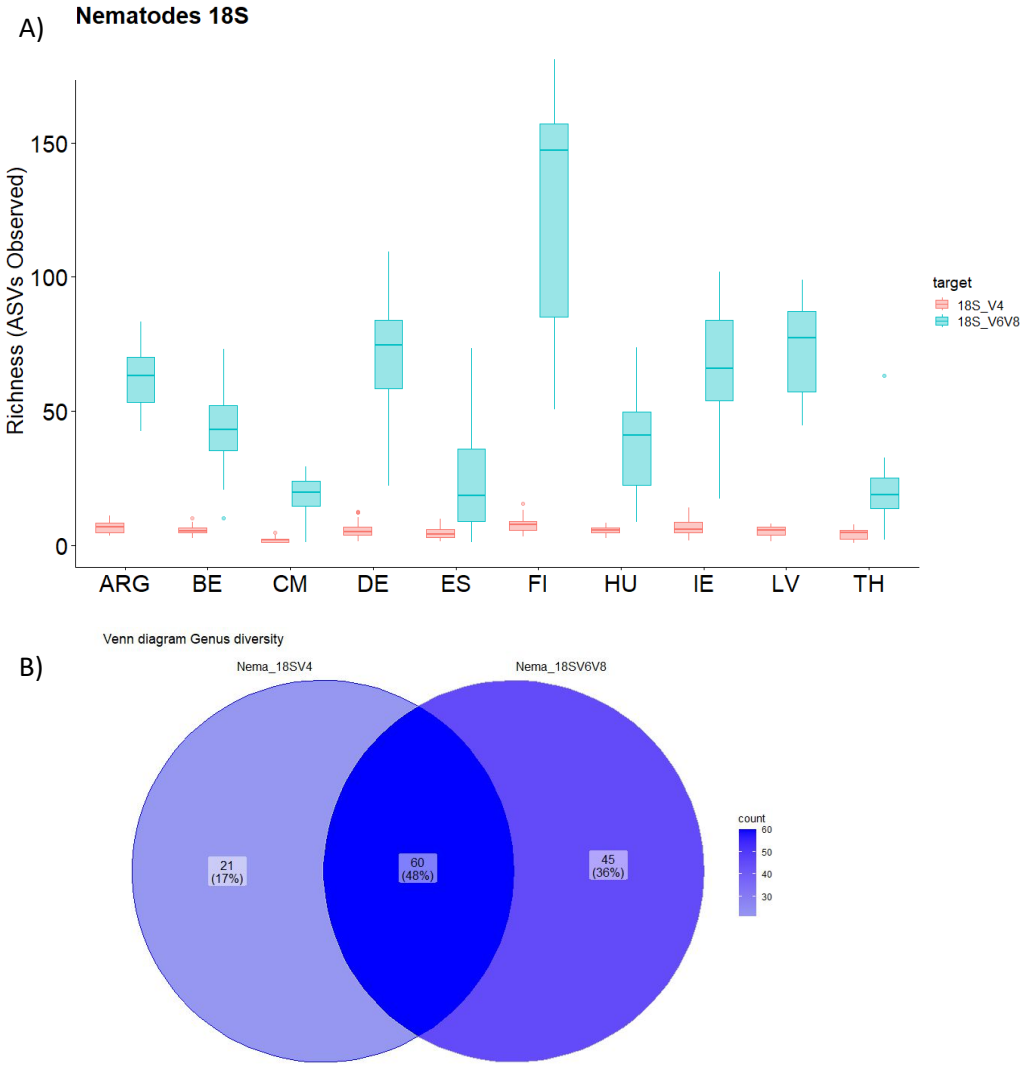


Figure S3. Alfa diversity distribution analysis for the different target regions/genes related to nematodes, divided by country. For the combination of country x target, the observed ASVs are represented by boxplots. B) Total of shared and non-shared genera between sequencing datasets.

Micro-arthropods (mites + collembola)

Micro-arthropods richness from 18SV6V8 is higher compared to the one from 18SV4, some for genus. Albeit we are aware of the limitations (waiting for the new primers sequencing), we select this sequencing dataset for further analysis related to micro-arthropods.

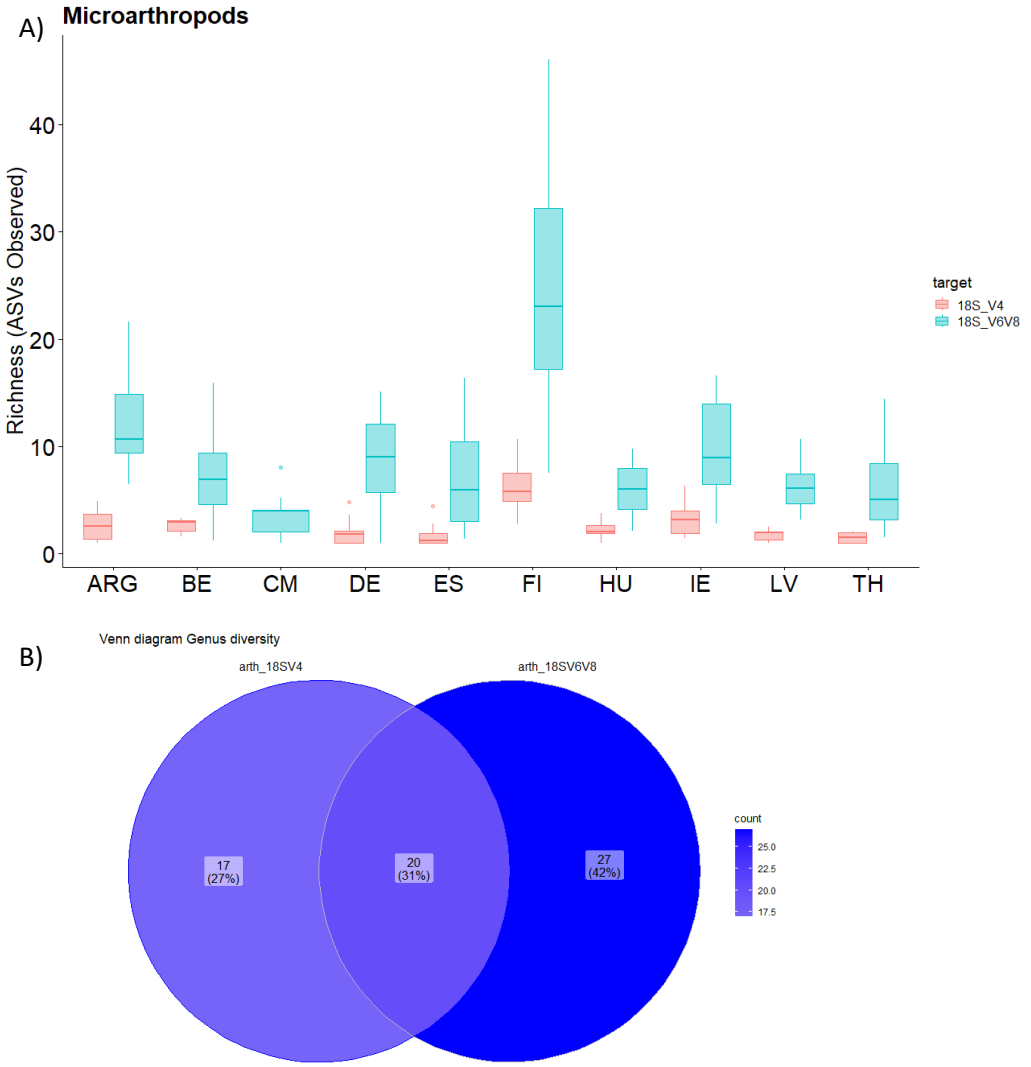
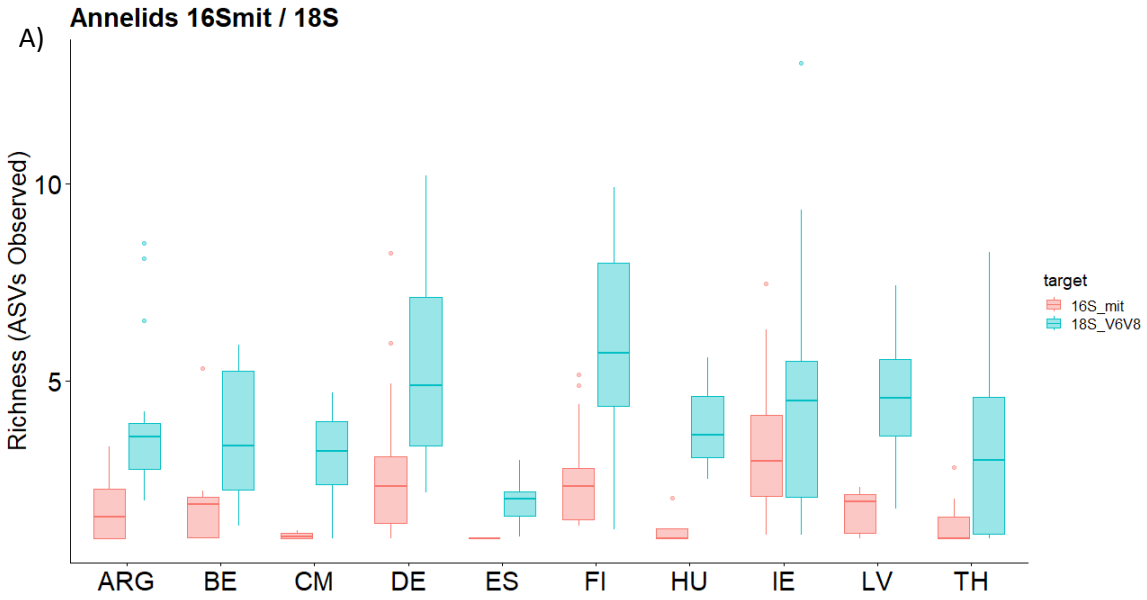


Figure S4. Alfa diversity distribution analysis for the different target regions/genes related to micro-arthropods (collembola + mites), divided by country. For the combination of country x target, the observed ASVs are represented by boxplots. B) Total of shared and non-shared genera between sequencing datasets.

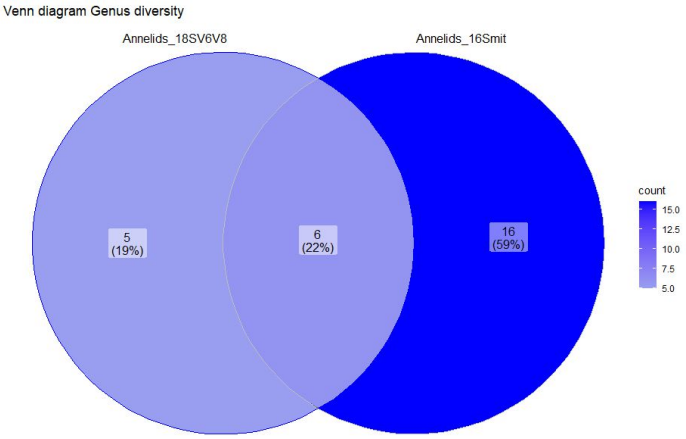
Annelids

Although the diversity observed in 18SV6V8 and the sample number retained are higher compared to 16Smit, the key point here is that we observe higher number genus and ASVs related to Lumbricidae (earthworms) family in 16Smit compared to 18SV6V8, with higher number of counts (>2,000). Thus, as earthworms is one of our soil interesting groups within annelids, we select this pair of primers to further analysis. In terms of 18SV4, we discarded it for now because after filtering the annelids, only 86 samples have more than 30 counts related to this group, encompassing 52 ASVs.

A)



B)



C)

18SV6V8	16Smit
Enchytraeidae (46)	Enchytraeidae (51)
Lumbricidae (5)	Lumbricidae (49)
Megascolecidae (1)	Megascolecidae (6)
Naididae (14)	Naididae (5)
Aeolosomatidae (1)	
Unclassified (101)	

Figure S5. *A) Alfa diversity distribution analysis for the different target regions/genes related to annelids, divided by country. For the combination of country x target, the observed ASVs are represented by boxplots. B) Total of shared and non-shared genera between sequencing datasets. C) All families identified in both sequencing datasets.*

Eukaryotes database curation

In this part, a "new" database has been created consisting of all PR2 v5.0 + SILVA v.138 (fungi and metazoa only), which allows a correct classification at different levels for all groups of eukaryotes, even increasing the diversity of families and genera classified with respect to original PR2. This taxonomy was harmonized between databases with taxonomizr v0.10.2 in R and using the classic Linnean taxonomic levels of the NCBI (accessed 20/02/2024) in the case of metazoans, or the levels of the UNITE v.8.3 database for Fungi. Since PR2 is a database mainly focused on protists, the classification and sequences of protists have been maintained, removing the protist information from the SILVA138 database prior to merging and dereplication. Dereplication and sequences curation was performed mainly with VSEARCH.

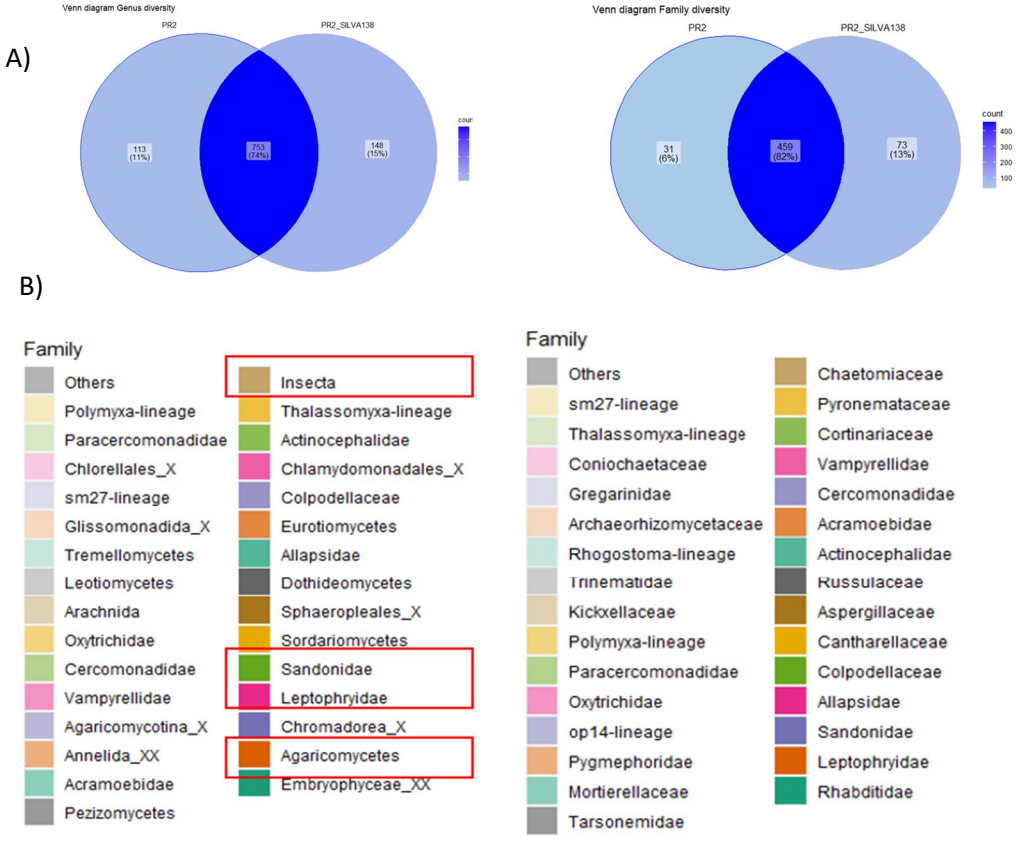


Figure S6. *A) Shared and no shared genus/families between both databases. B) Main families identified with PR2 databases with no modification (left) and main families with PR2 + SILVA138 harmonized (right).*

Table S3. *Multivariate effect of country (only EU croplands), management and soil degradation on beta-diversity after two-three way PERMANOVA analysis with 999 permutations.*

		prok		fungi		protist		nematodes		micro-arthropods		annelida	
	Factor	R2	p	R2	p	R2	p	R2	p	R2	p	R2	p
EU croplands	C	0.27	0.00	0.26	0.00	0.16	0.00	0.10	0.00	0.14	0.00	-	-
	M	0.01	0.08	0.02	0.00	0.01	0.02	0.02	0.00	0.02	0.08	-	-
	D	0.03	0.02	0.03	0.14	0.03	0.35	0.03	0.51	0.04	0.06	-	-
	CxM	0.04	0.04	0.05	0.00	0.05	0.01	0.05	0.00	0.06	0.05	-	-
	CxD	0.02	0.02	0.02	0.19	0.02	0.07	0.02	0.09	0.02	0.42	-	-
	MxD	0.02	0.23	0.02	0.49	0.03	0.39	0.03	0.14	0.03	0.93	-	-
	CxMxD	0.02	0.16	0.02	0.01	0.02	0.20	0.02	0.05	0.02	0.56	-	-
TH	M	0.05	0.06	0.04	0.24	0.04	0.04	0.03	0.78	0.06	0.23	-	-
	D	0.13	0.00	0.13	0.00	0.10	0.00	0.07	0.36	0.12	0.21	-	-
	MxD	0.09	0.06	0.10	0.00	0.09	0.01	0.10	0.03	0.11	0.31	-	-
ARG	M	0.04	0.76	0.07	0.04	0.06	0.05	0.05	0.50	0.07	0.08	-	-
	D	0.09	0.02	0.05	0.40	0.07	0.02	0.06	0.27	0.07	0.23	-	-
	MxD	0.04	0.78	0.04	0.83	0.05	0.51	0.04	0.90	0.04	0.75	-	-
CM	M	0.10	0.02	0.08	0.03	0.07	0.08	0.05	0.40	-	-	-	-
	D	0.06	0.30	0.07	0.07	0.06	0.34	0.06	0.28	-	-	-	-
	MxD	0.04	0.56	0.05	0.68	0.06	0.45	0.08	0.08				
Ire Grassland	M	0.03	0.95	0.05	0.47	0.05	0.62	0.07	0.10	0.04	0.67	0.02	0.95
	D	0.10	0.04	0.10	0.02	0.05	0.36	0.10	0.00	0.06	0.33	0.16	0.03
	MxD	0.03	0.95	0.04	0.92	0.04	0.95	0.06	0.20	0.05	0.48	0.02	0.97
Ire Agroforest	M	0.14	0.24	0.11	0.46	0.13	0.22	0.08	0.87	0.21	0.17	0.09	0.92
	D	0.10	0.58	0.10	0.68	0.10	0.87	0.08	0.86	0.16	0.53	0.14	0.45
	MxD	0.06	0.93	0.11	0.39	0.10	0.69	0.10	0.65	0.12	0.97	0.10	0.70
FI Forest	M	0.31	0.00	0.34	0.00	0.34	0.00	0.29	0.00	0.31	0.01	0.39	0.02

Table S4. *Variables used for biodiversity indicators.*

Variable	**Description**
nematodes_abundance	total number of nematodes per 100 g fw soil; site
nematodes_richness	number of species richness nematodes per 100 g fw soil

collembola_abundance	total number of collembola per 100 g fw soil site
collembola_richness	number of species richness collembola per 100 g
mites_abundance	total number of collembola per 100 g fw soil
mites_richness	number of species richness mites per 100 g fw soil per site
nag	nmol MUF·g soil-1·h-1
phos	nmol MUF·g soil-1·h-1
xyl	nmol MUF·g soil-1·h-1
bg	nmol MUF·g soil-1·h-1
leafdamage_fungi	Average percentage of leaf damage caused by pathogenic fungi
leafdamage_total	Average percentage of leaf damage caused by herbivores and pathogenic fungi
bacteria	(nmol PLFA/g soil)
fungi	(nmol PLFA/g soil)
amf	Arbuscular Mycorrhizal Fungi (nmol NLFA/g soil)
actinobacteria	(nmol PLFA/g soil)
methanotrophs	(nmol PLFA/g soil)
protists	(nmol PLFA/g soil)
microeukaryots	(nmol PLFA/g soil)
tmb	Total Microbial Biomass (nmol PLFA/g soil)
tmsb	Total Microbial Storage Biomass (nmol NFLA/g soil)
16s	Copies.ng DNA-1
aob	Copies.ng DNA-1
aoa	Copies.ng DNA-1
nir_k	Copies.ng DNA-1
nir_s	Copies.ng DNA-1
nos_z1	Copies.ng DNA-1
nos_z2	Copies.ng DNA-1
coma_a	Copies.ng DNA-1
coma_b	Copies.ng DNA-1

plant_parasitic_index	based no nematodes data
structure_index	% based no nematodes, informs about the level of available nutrients, information about how fast or slow goes the nutrient cycling
herbivore_footprint	based no nematodes data
root_herbivores_abundance	% of nematodes that are herbivores, multiplied by the amount of nematodes/100g
maturity_index	measure of soil disturbance based on nematodes, Bongers 90 oecologia
sobs_16S_iter	Alpha-diversity index (prokaryotes): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_16S_iter	Beta-diversity (prokaryotes): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_16S_iter	Beta-diversity (prokaryotes): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_18SV4_iter	Alpha-diversity index (global eukaryotes): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_18SV4_iter	Beta-diversity (global eukaryotes): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_18SV4_iter	Beta-diversity (global eukaryotes): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_ITS_iter	Alpha-diversity index (fungi): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_ITS_iter	Beta-diversity (fungi): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_ITS_iter	Beta-diversity (fungi): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_18SV4prot_iter	Alpha-diversity index (protist): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_18SV4prot_iter	Beta-diversity (protist): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_18SV4prot_iter	Beta-diversity (protist): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_18SV6V8_iter	Alpha-diversity index (global eukaryotes, mostly metazoa): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_18SV6V8_iter	Beta-diversity (global eukaryotes, mostly metazoa): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_18SV6V8_iter	Beta-diversity (global eukaryotes, mostly metazoa): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction

sobs_18SV6V8nema_iter	Alpha-diversity index (nematodes): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_18SV6V8nema_iter	Beta-diversity (nematodes): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_18SV6V8nema_iter	Beta-diversity (nematodes): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_18SV6V8arth_iter	Alpha-diversity index (micro-arthropods): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_18SV6V8arth_iter	Beta-diversity (micro-arthropods): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_18SV6V8arth_iter	Beta-diversity (micro-arthropods): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_18SV4V5_iter	Alpha-diversity index (global eukaryotes): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_18SV4V5_iter	Beta-diversity (global eukaryotes): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_18SV4V5_iter	Beta-diversity (global eukaryotes): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
sobs_16Smit_iter	Alpha-diversity index (annelids): richness estimation based on ASVs identified by sample, after rarefaction
Axis1_16Smit_iter	Beta-diversity (annelids): First principal component (PCO1) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
Axis2_16Smit_iter	Beta-diversity (annelids): Second principal component (PCO2) obtained after a Principal Coordinate analysis (PCoA) from Bray_Curtis dissimilarity matrix, after rarefaction
ploughing_depth	depth at which the ploughing is made
amount_chemical_fertilisation	kg/Ha
amount_organic_fertilisation	kg/Ha
ph	soil pH, measured in 1:5 soil_water extraction
clay	Clay content in soils (%)
conductivity	soil salinity (in micro-Siemens/cm), measured in 1:5 soil_water extraction
toc	Total Organic Carbon (g kg-1) as measured in the lab, following acid fumigation methodology
whc	% soil moisture (100% water holding capacity)
available_p	inorganic phosphates extracted by P Olsen method (mgP/kg soil)
amo	mg N-NH4 kg soil-1
din	mg N-NO3 kg soil-1

nit	mg N-(NO3+NH4) kg soil-1
tan	mg N-TAN kg soil-1
ecosystem_production	Based on NDVI mean provided by NOA
herbivore_footprint	based no nematodes data
enrichment_index	based no nematodes data
sand	Sand content in soils (%)
silt	Silt content in soils (%)
cstock	Carbon stock (Ton/ha), calculated as TOC*bulk_density*sampling depth (10cm)
litter_decomposition_rate_k	as measured with the tea bag index
stabilization_factor_s	related with tea bag index
bulk_density	Bulk density of dry soil (sieved) g/cm
infiltration	inverse of soil infiltration rate (laboratory procedure), seconds it takes to infiltrate 5 ml of water once saturated
amt	Annual Mean Temperature
mdr	Mean Diurnal Range (Mean of monthly (max temp - min temp))
iso	Isothermality (BIO2/BIO7) (* 100)
tse	Temperature Seasonality (standard deviation *100)
matwn	Max Temperature of Warmest Month
mitcm	Min Temperature of Coldest Month
atr	Annual Temperature Range (BIO5-BIO6)
mtweq	Mean Temperature of Wettest Quarter
mtdq	Mean Temperature of Driest Quarter
mtwaq	Mean Temperature of Warmest Quarter
mtcq	Mean Temperature of Coldest Quarter
ap	Annual Precipitation
pwm	Precipitation of Wettest Month
pdm	Precipitation of Driest Month
pse	Precipitation Seasonality (Coefficient of Variation)
pweq	ASV59286_16S est Quarter
pdq	Precipitation of Driest Quarter

pwaq	Precipitation of Warmest Quarter
pcq	Precipitation of Coldest Quarter

Annex IV. Detailed results for the permutational linear analyses performed to predict soil biodiversity and soil functioning

SOIL BIODIVERSITY

EUROPEAN NUTS-2 REGIONS

Table S1. *Summary of the linear models (estimate and R^2 adjusted) selected by AIC to assess the effect of different predictors on biodiversity groups in the European regions of the study (Belgium, Denmark, Hungary, Latvia, and Spain, n=110). Positive effects are represented in green and negative effects in red. Predictors highlighted in gray were not considered in model construction. Significance codes: * < 0.05, ** < 0.01, *** < 0.001. Data was imputed for NA values when they were <33% for the variable. Abbreviations → NEI: Naturalness evaluation index; SWF: Small wood features; NEM: Richness of nematodes; COLLEMB: Richness of collembolan; MITES: Richness of mites; BACT: Bacteria; AMF: Arbuscular mycorrhizal fungi; ACTINOB: Actinobacteria; METHAN: Methanotrophs; PROTIST: Protists; MICROEUK: Microeukaryots; MESO: Mesofauna; PLFA_NLFA: Microbial community (plfa and nfla), ROOT H: Root herbivores (nematodes) abundance.*

	NEMAT	COLLEMB	MITES	BACT	FUNGI	AMF	ACTINOB	METHAN	PROTIST	MICROEUK	ROOT_H
SOIL DEGRADATION (SD)	-0.14		3.035e-03	0.12	0.13	-0.03	0.007	-0.003		-0.1	-0.2
SOIL MANAGEMENT (MAN)	-0.08	0.24*	7.728e-02	-0.4*	-0.33	-0.14	-0.44	-0.36		0.17	-1.83
ARIDITY (AR)	-0.61		-2.308e-01	0.62	0.7	-0.05	0.3	0.23		-0.39	
NEI				-0.73	-1*	0.12	-0.47	-0.52			
SWF			-9.318e-06	1.828e-05	1.979e-05		2.656e-05	0			8.829e-07
SAND	0.16			-2.147			-0.26*	-0.1		0.05	0.096
BULK DENSITY	-0.07				-0.25*						-0.17
PH	-0.02				-0.011	0.21*	0.07			0.065	-0.11
SD X MAN	0.34		3.856e-01	-0.35	-0.05	-0.6*	-0.41	0.011			
SD X AR	-1.06		-7.018e-01	0.14	-0.06	-0.01	-0.008	0.24		-0.6	
MAN X AR	0.74		-1.715e-01	-0.19	-0.6	-1.7	-0.23				
MODEL R^2 ADJUSTED	**0**	**0**	**0**	**0.03**	**0.02**	**0.02**	**0.05**	**0**		**-0.04**	**0.03**

	ITS	16S	18S	BIODIVERSITY
SOIL DEGRADATION (SD)	-0.25	-0.33	-0.12	-0.15
SOIL MANAGEMENT (MAN)	-0.36	-0.06	-0.45*	-0.12
ARIDITY (AR)	-0.48	-1.11	-0.26	-0.64*
NEI	-0.75	-0.43		
SWF		-8.630e-06	0	
SAND	0.28			
BULK DENSITY	-0.07	0.17		0.03
PH	0.02	0.08	-0.07	-0.02
SD X MAN	0.13	-0.01		0.22*
SD X AR	-0.4	-0.01*	0.15	-0.65*
MAN X AR	-0.44	-1.01	-0.43	
MODEL R^2 ADJUSTED	**0.1**	**0.03**	**0.03**	**0.04**

EUROPEAN NUTS-2 AND INTERNATIONAL REGIONS

Table S2. *Summary of the linear models (estimate and R^2 adjusted) selected by AIC to assess the effect of different predictors on biodiversity groups in the European (Belgium, Denmark, Hungary, Latvia, and Spain) and international regions (Argentina, Cameroon and Thailand) of the study (n=180). Positive effects are represented in green and negative effects in red. Predictors highlighted in gray were not considered in model construction. Significance codes: * < 0.05, ** < 0.01, *** < 0.001. Data was imputed for NA values when they were <33% for the variable. Abbreviations → NEI: Naturalness evaluation index; SWF: Small wood features; NEM: Richness of nematodes; COLLEMB: Richness of collembolan; MITES: Richness of mites; BACT: Bacteria; AMF: Arbuscular mycorrhizal fungi; ACTINOB: Actinobacteria; METHAN: Methanotrophs; PROTIST: Protists; MICROEUK: Microeukaryots.*

	ITS	16S	18S	BACT	FUNGI	AMF	ACTINOB	METHAN	PROTIST	MICROEU	BIODIVERSITY
SOIL DEGRADATION (SD)	-0.16	-0.24*	-0.12	0.064	-0.023	-0.08	0.023	0.002		-0.06	-0.13*
SOIL MANAGEMENT (MAN)	-0.24*	-0.056	-0.36**	-0.26	-0.21	0.13	-0.21	-0.24*		-0.11	-0.14
ARIDITY (AR)	-0.2	-0.42	-0.12	0.048	0.17	0.15	-0.07	0.18		-0.17	-0.17
NEI	-0.22	-0.36	-0.11	-0.016	-0.55	-0.01	-0.11				-0.24
SWF		1.668e-05	2.691e-05	3.595e-05	3.598e-05	0	3.624e-05				2.153e-05
SAND	0.17*	1.41*		-0.22***		0.15	-0.23				
BULK DENSITY		0.18	0.064		-0.25***	-0.1	-0.23**			-0.02	
PH			0.026	-0.04	0.046	0.19	-0.053	0.17*		0.03	0.042
SD X MAN	0.15	0.28	-0.07	-0.17	-0.028			-0.09		-0.002	0.25*
SD X AR	-0.03	-0.45	0.26	-0.039	0.22	-0.36	-0.28	0.25		-0.34	0.25
MAN X AR	0.11	0.14	-0.13	0.67	0.36	-0.77	0.61			-0.34	0.28
MODEL R^2 ADJUSTED	**0.03**	**0.05**	**0.02**	**0.06**	**0.08**	**0.03**	**0.1**	**0.01**		**0**	**0.04**

Table S3. *Six metrics as calculate for each of 23 co-occurrence networks: number of nodes, number of edges, connectance, modularity, transitivity and assortativity. Each metric is followed by standard error ($p < 0.05$) as calculated based on a null model distribution for that metric, based on 1000 permutations of each observed network. Bold characters for modularity, transitivity and assortativity indicate values that are significantly different from the null model distribution ($p < 0.05$), and therefore represent network properties can be attributed to biological patterns rather than to chance.*

NUTS-2 region	Management	Nodes (nr)	Edges (nr)	Connectance	Modularity	Transitivity	Assortativity
West Flanders	ALT	1226	7251 ± 42	0.01 ± 0.0001	**0.914** ± 0.0018	0.999 ± 0.0003	**0.089** ± 0.004
West Flanders	CON	1234	7375 ± 43	0.01 ± 0.0001	**0.874** ± 0.0019	0.983 ± 0.0003	**0.062** ± 0.004
Murcia	ALT	1476	60967 ± 125	0.056 ± 0.0001	**0.395** ± 0.0009	0.996 ± 0.0001	**0.869** ± 0.001
Murcia	CON	1532	71667 ± 127	0.061 ± 0.0001	**0.496** ± 0.0008	1 ± 0.0001	**0.834** ± 0.001
Latvia	ALT	1421	9132 ± 47	0.009 ± 0	**0.883** ± 0.002	0.988 ± 0.0002	**0.075** ± 0.004
Latvia	CON	1468	13154 ± 58	0.012 ± 0.0001	**0.698** ± 0.0029	0.991 ± 0.0002	**0.059** ± 0.003
South Transdanubia	ALT	2721	599407 ± 346	0.162 ± 0.0001	**0.037** ± 0.0003	1 ± 0.0001	**0.043** ± 0
South Transdanubia	CON	1501	11510 ± 54	0.01 ± 0	**0.904** ± 0.0028	0.919 ± 0.0002	**0.158** ± 0.003
Middle Jutland/South Denmark	ALT OLD	1549	10961 ± 52	0.009 ± 0	**0.88** ± 0.0023	0.994 ± 0.0002	**0.064** ± 0.003
Middle Jutland/South Denmark	ALT YOUNG	1344	8185 ± 45	0.009 ± 0	**0.899** ± 0.0018	0.997 ± 0.0003	**0.061** ± 0.004
Middle Jutland/South Denmark	CON	1344	9877 ± 48	0.011 ± 0.0001	**0.848** ± 0.0028	0.996 ± 0.0002	**0.025** ± 0.003
Buenos Aires	ALT	1160	6680 ± 41	0.01 ± 0.0001	**0.842** ± 0.002	0.997 ± 0.0003	**0.058** ± 0.004
Buenos Aires	CON	1381	12406 ± 55	0.013 ± 0.0001	**0.619** ± 0.0028	0.996 ± 0.0002	**0.08** ± 0.003
West Cameroon	ALT	1125	6894 ± 41	0.011 ± 0.0001	**0.884** ± 0.0022	0.983 ± 0.0003	**0.147** ± 0.005
West Cameroon	CON	2331	449556 ± 295	0.166 ± 0.0001	**0.156** ± 0.0004	1 ± 0.0001	**0.881** ± 0
Chiangrai	ALT	631	4458 ± 32	0.022 ± 0.0002	**0.87** ± 0.0024	1 ± 0.0005	**0.145** ± 0.005
Chiangrai	CON	789	10335 ± 50	0.033 ± 0.0002	**0.64** ± 0.0018	1 ± 0.0003	**0.129** ± 0.004

Southern Ireland	*MIXTURE*	*2268*	*440365*	*±*	*295*	*0.171*	*±*	*0.0001*	***0.021***	*±*	*0.0003*	*1*	*±*	*0.0001*	***0.131***	*±*	*0.001*
Southern Ireland	*MONOCULTURE*	*1252*	*6947*	*±*	*42*	*0.009*	*±*	*0.0001*	***0.87***	*±*	*0.0017*	*0.996*	*±*	*0.0003*	***0.031***	*±*	*0.004*
Southern Ireland	*AGROFORESTRY*	*1768*	*356287*	*±*	*260*	*0.228*	*±*	*0.0002*	***0.022***	*±*	*0.0003*	*1*	*±*	*0.0002*	***0.184***	*±*	*0.001*
Southern Ireland	*AG. CONTROL*	*797*	*4926*	*±*	*35*	*0.016*	*±*	*0.0001*	***0.905***	*±*	*0.0028*	*0.996*	*±*	*0.0004*	***0.074***	*±*	*0.005*
Western Finland	*CONT. COVER*	*1955*	*548903*	*±*	*317*	*0.287*	*±*	*0.0002*	***0.173***	*±*	*0.0003*	*1*	*±*	*0.0002*	***0.969***	*±*	*0*
Western Finland	*CLEAR CUTTING*	*1908*	*923073*	*±*	*329*	*0.507*	*±*	*0.0002*	***0.004***	*±*	*0.0002*	*1*	*±*	*0.0002*	***0.001***	*±*	*0*

SOIL FUNCTIONS

EUROPEAN NUTS-2 REGIONS

Table S4. *Summary of the linear models (estimate and R^2 adjusted) selected by AIC to assess the effect of different predictors on soil functions in the European NUTS-2 regions (Belgium, Denmark, Hungary, Latvia, and Spain) of the study (n=110). Positive effects are represented in green and negative effects in red. Predictors highlighted in gray were not considered in model construction. Significance codes: * < 0.05, ** < 0.01, *** < 0.001. Data was imputed for NA values when they were <33% for the variable. The biodiversity variable was constructed as the standardized average of ITS, 18S, 16S and richness of nematodes, mites and collembolla. Abbreviations → NEI: Naturalness evaluation index; SWF: Small wood features; TOC: Total organic carbon; AVAP: Available P; TAN: Total Available N; AMO: Ammonification; DEP: Depolymeratization; NTR: Potential nitrification rate; BG: beta-glucosidase; XYL: xylanase; PHOS: phosphatase; NAG: N-acetylglucosaminidase; CONDUC: Conductivity; WHC: Water holding capacity; INFILT: Infiltration; AGGR: Aggregates (Soil erosion resistance); LD: Litter decomposition rate; LDF: Leaf damaged caused by fungi; LDH: Leaf damage caused by herbivores (pathogens); N RET: N retained in soil (fertilizer); P RET: P retained in soil (fertilizer); MF: Multifunctionality index.*

	TOC	AVAP	TAN	AMO	DEP	NTR	BG	XYL	PHOS	NAG
SOIL DEGRADATION (SD)	0.056	0.15	0.29*	0.09	-0.18	-0.19	0.37*	0.32**	0.13	0.19
MANAGEMENT (MAN)	-0.35*	-0.11	-0.16	-0.13	-0.09	-0.26	-0.16	-0.49*	-0.18	-0.038
ARIDITY (AR)	-0.031	-0.32	0.96*	0.31	-0.59	-0.61	1.69**	1.31**	0.32	0.45
NEI	0.35		-0.3	0.21	0.87*	0.17	-0.83		-0.22	0.034
SWF	9.857e-06			5.157e-05	-7.763e-06	2.062e-05	-1.517e-05		-1.084e-05	8.175e-06
SAND	-0.33***	-0.14	-0.42***	-0.31**	-0.36***	-0.49***			0.1	-0.11
BULK DENSITY	-0.24*	0.003	-0.2*		-0.002			-0.35***	-0.029	0.006
PH	0.036	0.07	-0.02		0.35***	0.14*	0.19*	0.005		-0.088
BIODIVERSITY (BIO)	0.59***	0.1	0.19	0.49*	0.13	0.55***	0.11	0.46**	0.56***	0.25
SD X MAN			-0.31	0.17	0.38	-0.1	-0.16	-0.21	0.09	-0.18
SD X AR	-0.5	-1.08	0.8	1.9	0.26	-0.55	1.17	1.28*	-0.07	
MAN X AR	0.3	1.3	0.85	1.9*	1.05	0.85	-0.31		0.72	-0.5
BIO X MAN	0.5	0.46	0.83	1.3***	-0.14				0.43	
BIO X SD	0.1	-0.4	-0.46	-0.58*					0.03	-0.14
BIO X AR	0.54	-1.44	-0.29	-0.86	-0.96		1.23*	0.65	0.48	-0.28
MODEL R^2 ADJUSTED	**0.21**	**0.02**	**0.13**	**0.12**	**0.29**	**0.34**	**0.07**	**0.19**	**0.03**	**-0.07**

	CONDUC	WHC	INFILT	AGGR	LD	LDF	LDH	N RET	P RET	MF
SOIL DEGRADATION (SD)	0.07	-0.033	-0.15	0.11	-0.18	-0.043	0.04	0.07	-0.15	0.013
SOIL MANAGEMENT (MAN)	-0.06	-0.025	-0.15	-0.68***	-0.53**	0.16	0.19	0.25	-0.11	-0.15*
ARIDITY (AR)	0.3	0.16	-0.53	0.46	-0.07	0.016	0.22	0.46	-0.32	0.09
NEI			0.064			-0.094			0.66	
SWF		8.963e-06	-9.176e-06	4.777e-06		-1.289e-06		2.148e-05	1.222e-05	7.060e-08
SAND		-0.53***		0.11	-0.004			0.044	-0.45***	-0.10
BULK DENSITY	-0.008	-0.017	-0.14	-0.16	0.08	-0.039	-0.05	0.033	-0.11	0.016
PH	0.2*	0.17		-0.026	0.07	0.13	-0.03	-0.1		0.052
BIODIVERSITY	-0.41	0.18	0.074	0.31	-0.15	-0.066	0.24	-0.065	0.13	0.23***
SD X MAN	0.38	0.2	-0.57	-0.35		-0.15	0.22	0.04	-0.2	-0.051
SD X AR		5.037e-03	-0.15		0.22	0.52		0.64	-4.835e-03	
MAN X AR		0.54	-1.42			-0.84		-0.49	-0.9	
BIO X MAN	0.24	0.012		0.78*		0.09	-0.45	-0.17	-0.034	0.11
BIO X SD	-0.23	0.57*		-0.2		0.25	0.28	-0.34	0.43*	-0.07
BIO X AR		0.85		0.13	0.52	0.79		-0.41	1.09	-0.16
MODEL R^2 ADJUSTED	**0.14**	**0.37**	**0**	**0.19**	**0.05**	**-0.03**	**0.04**	**-0.06**	**0.25**	**0.12**

EUROPEAN NUTS-2 AND INTERNATIONAL REGIONS

Table S6. *Summary of the linear models (estimate and R^2 adjusted) selected by AIC to assess the effect of different predictors on soil functions in the European NUTS-2 (Belgium, Denmark, Hungary, Latvia, and Spain) and international regions (Argentina, Cameroon and Thailand) of the study (n=180). Positive effects are represented in green and negative effects in red. Predictors highlighted in gray were not considered in model construction. Significance codes: * < 0.05, ** < 0.01, *** < 0.001. Data was imputed for NA values when they were <33% for the variable. The biodiversity variable was constructed as the standardized average of ITS, 18S and 16S. Abbreviations → NEI: Naturalness evaluation index; SWF: Small wood features; TOC: Total organic carbon; AVAP: Available P; TAN: Total Available N; AMO: Ammonification; DEP: Depolymeratization; NTR: Potential nitrification rate; BG: beta-glucosidase; XYL: xylanase; PHOS: phosphatase; NAG: N-acetylglucosaminidase; CONDUC: Conductivity; WHC: Water holding capacity; INFILT: Infiltration; AGGR: Aggregates (Soil erosion resistance); LD: Litter decomposition rate; LDF: Leaf damaged caused by fungi; LDH: Leaf damage caused by herbivores (pathogens); N RET: N retained in soil (fertilizer); P RET: P retained in soil (fertilizer); MF: Multifunctionality index.*

	CONDUC	WHC	INFILT	AGGR	LD	LDF	LDH	N RET	P RET	MF
SOIL DEGRADATION (SD)	-1.746e-03	-0.12	-0.08	-0.11	-0.18**	-0.02	0.08	-0.036	-0.1	-3.687e-03
SOIL MANAGEMENT (MAN)	-0.06	-4.924e-03	0.11	-0.21	-0.23	0.21	0.17	0.31	-0.05	-0.07
ARIDITY (AR)	0.24	-0.41	-0.18	0.06	-0.21	-0.39	-0.53	0.12	-0.25	0.037
NEI	-0.09	0.4	-0.19	-0.85***	0.17		0.19	-0.52	0.16	
SWF	2.535e-05	1.612e-05	-8.898e-06	-2.487e-06	1.259e-05		1.479e-05	3.949e-06	-0.34	-3.858e-06
SAND	-0.26**		0.13	0.16	0.04		-0.14*	0.045	-0.34***	-0.08***
BULK DENSITY	0.07	-0.15*	-0.12	-0.13	0.14*	-0.03	0.04	0.056		-6.173e-04
PH	0.21*	0.27***	-0.04	-0.063		0.16**	-0.06	-0.095	0.25***	0.05
BIODIVERSITY	-0.16	-0.11	0.04	0.23*	0.053	0.01	0.05	0.033	0.03	0.15***
SD X MAN	0.11	0.11	-0.51*		-0.23	-0.1	0.29	0.12	-0.004	-0.13*
SD X AR	0.34	-0.79*	-0.21	-0.11	0.11		-0.51		-0.3	3.729e-03
MAN X AR	0.06	0.26	-1.12	-0.62	0.5		0.31	-0.23	-0.15	0.1
BIO X MAN	0.09	0.16		0.45			0.044	-0.22	0.36	
BIO X SD		0.34***	-0.011	-0.034				-0.18	0.29	
BIO X AR	0.37	-3.648e-04	-0.13		0.44	0.22	0.47	-0.4	1.08	0.17
MODEL R^2 ADJUSTED	**0.08**	**0.11**	**0.01**	**0.08**	**0.06**	**0.03**	**0.03**	**-0.02**	**0.23**	**0.1**

	TOC	AVAP	TAN	AMO	DEP	NTR	BG	XYL	PHOS	NAG
SOIL DEGRADATION (SD)	0.05	-6.997e-03	0.11	-0.06	-0.08	-0.15	0.21*	0.25*	0.05	0.055
SOIL MANAGEMENT (MAN)	-0.31*	0.11	6.344e-03	2.256e-03	0.03	-0.11	-0.17	-0.53***	-0.23	0.04
ARIDITY (AR)	-0.16	-0.29	0.22	-0.13	-0.04	-0.03	0.99***	0.29	0.036	0.74*
NEI	-0.2		-0.32	-0.09	0.06	-0.37	-0.5	0.49	-0.08	-0.31
SWF	1.045e-05	1.093e-05	1.292e-05	2.086e-05	-3.205e-05	1.031e-05	-1.006e-05	-1.938e-05	-1.137e-06	2.259e-06
SAND	-0.25**	-0.09	-0.27***	-0.2	-0.2***	-0.4***	-0.14	-0.25***	-0.06	-0.11
BULK DENSITY	-0.3***	0.04	-0.15	-0.018	-3.564e-03		-0.13			-0.1
PH	0.08	0.09	-0.025			0.08	0.17		-0.4***	-0.023
BIODIVERSITY	0.3**	0.16	0.19	0.22	-0.11	0.25*	0.23	0.5***	0.31***	0.21
SD X MAN		-0.2	-0.47*	-0.48	0.27	8.278e-03	-0.07	-0.15	-2.370e-03	-0.18
SD X AR	-0.25	-0.35	0.33	0.07	-0.03	-0.42	1.4***	0.18	-0.22	1.3***
MAN X AR	0.38	1.25*	0.98	0.43	0.78	0.9	0.31	0.15	0.6	0.11
BIO X MAN	0.26	0.26	0.06	0.31	-0.08	-0.02	-0.06		-0.14	
BIO X SD	0.13	-0.05		-0.08	-0.1	0.12			-0.08	0.027
BIO X AR	0.62			-0.14	0.18	0.25	0.71	0.5	0.48	0.51
MODEL R^2 ADJUSTED	**0.18**	**0.03**	**0.11**	**0.06**	**0.02**	**0.2**	**0.08**	**0.12**	**0.15**	**0.007**

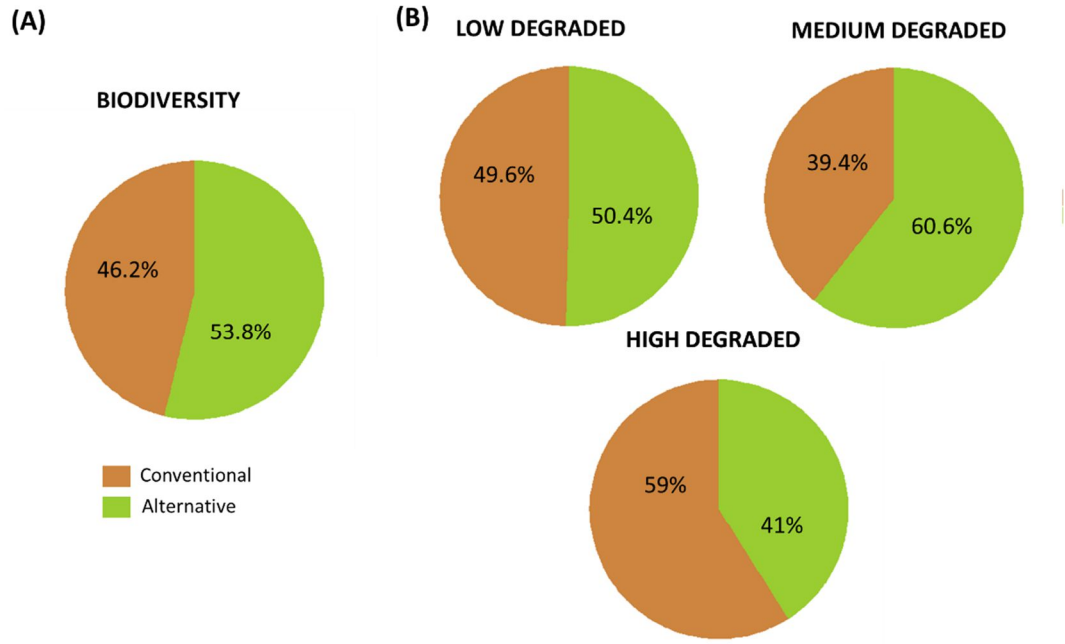


***Figure S1**. Simulated landscapes (using observed field data) maximizing soil biodiversity (A) at the landscape (5 pooled sites) scale and considering soil degradation levels (low, medium and high degraded) (B). Different colours show the proportion of conventional (brown) vs alternative (green) soil management required to maximize soil biodiversity across all our sites. To obtain the highest values of biodiversity across organisms and sites, we used a biodiversity index averaging the standardized values for each soil organism considered.*

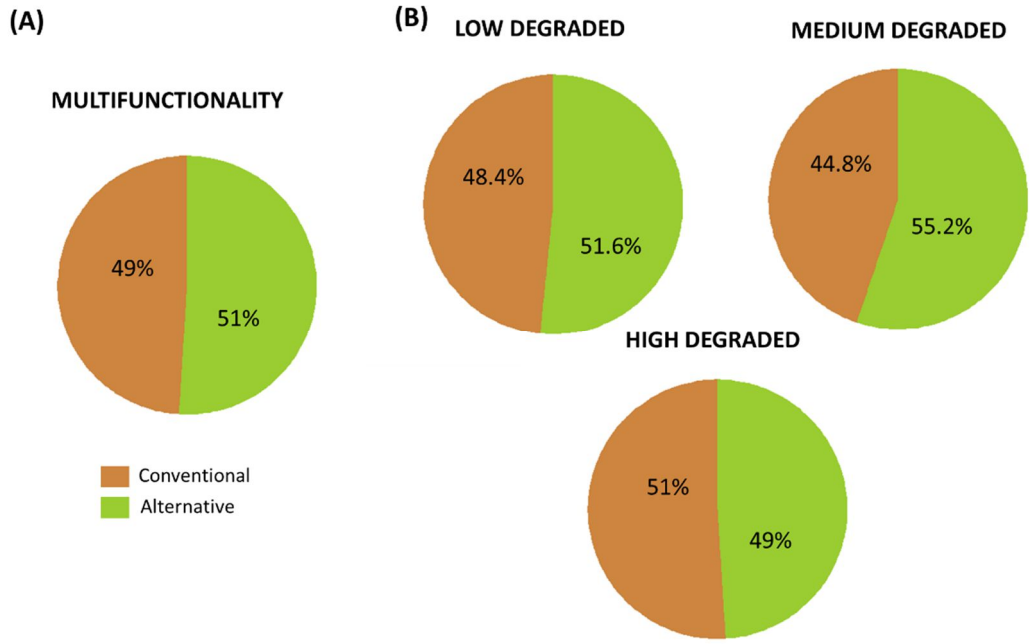


***Figure S2**. Simulated landscapes (using observed field data) maximizing soil multifunctionality (A) at the landscape (5 pooled sites) scale and considering soil degradation levels (low, medium and high degraded) (B). Different colours show the proportion of conventional (brown) vs alternative (green) soil management required to maximize soil multifunctionality across all our sites. To obtain the highest values of multifunctionality across sites, we used a multifunctionality index averaging the standardized values for each soil function considered.*

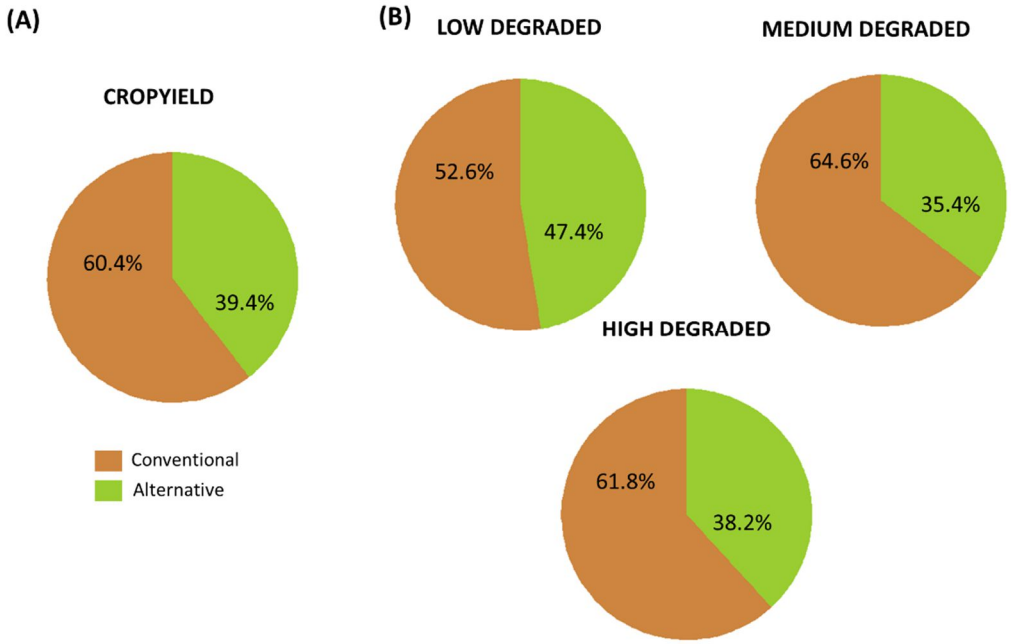


***Figure S3**. Simulated landscapes (using observed field data) maximizing agriculture crop yield (A) at the landscape (5 pooled sites) scale and considering soil degradation levels (low medium and high degraded) (B). Different colours show the proportion of conventional (brown) vs alternative (green) soil management required to maximize crop yield across all our sites. To achieve the highest crop yields across the sites, we measured crop production in kilograms per hectare per year.*

Annex V. Introducing the Sentinel-2 dataset & the Remote Sensing Indices

1. Introduction

The initial available imagery data from the polar-orbiting Sentinel-2 (S2) constellation can be traced back to 2015 and has proven to be a valuable resource across a wide spectrum of services and applications within the Copernicus Program [1]. These applications span various domains such as land monitoring, climate change analysis, emergency management, and more.

In addition, one of the key attributes of Sentinel-2 data is their open and free accessibility, which has played a pivotal role in democratizing access to high-quality Earth observation data. This accessibility has empowered global research and diverse applications.

Furthermore, the imagery data from the Sentinel-2 constellation boasts a temporal resolution of 5 days, as measured by the combined revisit frequency at the equator. Furthermore, it offers optical data with the finest spatial resolution available at no cost (10 meters), rendering it exceptionally suitable for meticulously monitoring dynamic environmental changes [1].

In particular, this dataset encompasses several spectral bands that capture distinct portions of the electromagnetic spectrum, including the blue, green, red, near-infrared, and short-wave infrared regions. Researchers and scientists worldwide harness this wealth of information for a multitude of purposes. Specifically, they employ equations, algorithms, and machine learning/artificial intelligence (ML/AI) models to extract meaningful insights from these data, a practice that holds immense importance for a broad array of scientific disciplines, including Remote Sensing (RS), Ecosystem Services (ES), Environmental Science, Agriculture, Forestry, and Geospatial Analysis [2-6], among others.

Recognizing the capabilities and potential applications of RS data and geoinformatics, the SOILGUARD consortium made the strategic decision to incorporate these technologies into the project. In particular, for the SOILGUARD project's specific needs, the Normalized Differential Vegetation Index (NDVI) and its statistical information, as well as the Ecosystem Stability Index (ESI), were acquired based on the methodology outlined in the study conducted by Garcia-Palacios et al. in 2018 [3].

More precisely, the NDVI was employed as a surrogate indicator for aboveground biomass [7-9], given its ability to provide a global assessment of vegetation health or "greenness" [10]. As for the ESI index, it served as a metric to gauge the stability of the ecosystems under examination and their capacity to deliver ecosystem services [3].

Sentinel-2 Data & Remote Sensing Indices

For acquiring information about the Remote Sensing indices of the project's NUTS-2(or equivalent) regions, the S2 dataset was downloaded and processed via the Google Earth Engine API (Table 1).

Table 23. Materials used and relevant information.

Dataset	Coverage	Spatial Resolution (m)	Temporal Reference	Source / Manual
Sentinel-2 constellation	Global	10	2018 - 2022	GEE / S2 Manual

2. Material & Methods

The initial phase in generating the Remote Sensing (RS) indices involves pre-processing the Sentinel-2 (S2) data within the Google Earth Engine (GEE) API. To initiate this process, the S2 Level-2A product available in the GEE data repository, was utilized. This product is known for its atmospheric correction, including cirrus clouds correction, ensuring that it accurately represents surface reflectance values [11]. The selection of this product not only guarantees the use of high-quality data but also reduces computation time.

Once the image collection of S2 data was identified, the dataset was further refined the by implementing a cloud mask [12] and setting a cloud percentage threshold, ensuring that the cloud cover remained below 20-30%. Following this pre-processing step, the spectral bands were utilized to derive the RS indices, as outlined in the study conducted by Garcia-Palacios et al. in 2018 [3], employing the capabilities of the GEE.

In particular, the RS indices were computed over a span of 5 years, specifically during the cropping period as detailed in Table 24. Additionally, calculations were performed for various buffer options surrounding the sampling locations. Specifically, the Normalized Differential Vegetation Index (NDVI) and its statistical information (mean and standard deviation) were computed for a buffer with a diameter of 200 meters around the sample locations, while the Ecosystem Stability Index (ESI) was determined for buffers of 200, 500, and 1000 meters in diameter encompassing the sampling points (an example can be seen in Figure 3).

Table 24. Cropping periods per country of the SOILGUARD Project.

Country	Cropping Period	Duration (months)	Temporal Reference
Belgium	April - November	8	2018 - 2022
Spain	May - July	3	2018 - 2022
Finland	May - August	4	2018 - 2022
Latvia	April - October	7	2018 - 2022
Hungary	April - October	7	2018 - 2022
Denmark	April - August	5	2018 - 2022
Ireland	March -November	9	2018 - 2022
Thailand	May - December	8	2018 - 2022
Argentina	June - December	7	2018 - 2022
Cameroon	April - July	4	2018 - 2022

Figure 3. Mean NDVI value (green hue represents higher values) of three example sample locations during December of 2022, Northern Thailand (via GEE).

3. Results

To ensure the meaningful and error-free interpretation and utilization of the results, it is essential to consider several key aspects in advance. These aspects pertain to:

1. **Interpreting NDVI Values:** NDVI values range from -1 to 1. Negative values typically indicate the presence of clouds and/or water, while values near zero often signify bare soil. Higher positive values (ranging from 0.1 to 0.5) suggest sparse vegetation, while values exceeding 0.6 indicate denser vegetation [7-10, 13].
2. **Interpreting ESI Values:** ESI values can be both negative and positive. Negative values indicate disturbances in the ecosystem stability or may represent other land use categories, such as lakes. Conversely, positive values indicate higher ecosystem stability, with higher values signifying greater stability [3].
3. **Handling "Clouds" Cells:** In some cases, S2 data may be missing due to factors like cloud cover, cloud shadows, technical issues during satellite image acquisition, and more. In such instances, these issues result in No Data values, which cannot be incorporated into data analysis. These cases are represented by the string 'Clouds' in the Excel file.
4. **Caution in Data Usage:** There are situations where, despite the presence of a mean NDVI value for the examined point, buffer, and time period, the Standard Deviation and the ESI have no information and are labeled as "Clouds." In such cases, it is advisable to exercise caution when using the mean value since it may solely represent one or a few image captures with data for the given month. If used, it should be noted that the accuracy of such data is limited.
5. **Consideration of Spatial and Temporal Resolution:** When analyzing and utilizing the data, it's important to take into account the spatial and temporal resolution. Specifically, the 10-meter spatial resolution implies that within a 100 square meter surface area, the spectral information or spectral index value represents the mean value of all elements within that surface, thus involving some degree of information aggregation. This limitation in terms of spatial resolution, although the finest available without charge, should be acknowledged. As for the temporal resolution, the 5-day revisit time of S2 data (at the equator) indicates that, theoretically, there could be up to 5-6 image acquisitions per month for the examined study site. However, due to factors like cloud cover, this number may be significantly lower in practice.

Remote Sensing Indices across EU NUTS-2 regions

All the results can be found in the provided excel file.

Remote Sensing Indices across International equivalent to NUTS-2 regions

All the results can be found in the provided excel file. It is noteworthy that there were no available S2 data before December of 2018 for the international sites of interest.

4. References

[1] Sentinel-2 - Missions - Sentinel Online - Sentinel Online. (n.d.). Sentinel Online. https://sentinel.esa.int/web/sentinel/missions/sentinel-2

[2] García-Palacios, P., Gross, N., Gaitán, J., & Maestre, F. T. (2018). Climate mediates the biodiversity–ecosystem stability relationship globally. Proceedings of the National Academy of Sciences of the United States of America, 115(33), 8400–8405. https://doi.org/10.1073/pnas.1800425115

[3] Furberg, D., Ban, Y., & Nascetti, A. (2019). Monitoring of Urbanization and Analysis of Environmental Impact in Stockholm with Sentinel-2A and SPOT-5 Multispectral Data. Remote Sensing, 11(20), 2408. https://doi.org/10.3390/rs11202408

[4] Cazorla, B. P., Cabello, J., Reyes, A., Guirado, E., Peñas, J., Pérez-Luque, A. J., & Alcaraz-Segura, D. (2023, April 27). A remote-sensing-based dataset to characterize the ecosystem functioning and functional diversity in the Biosphere Reserve of the Sierra Nevada (southeastern Spain). Earth System Science Data. https://doi.org/10.5194/essd-15-1871-2023

[5] Francini, M., Salvo, C., Viscomi, A., & Vitale, A. (2022). A Deep Learning-Based Method for the Semi-Automatic Identification of Built-Up Areas within Risk Zones Using Aerial Imagery and Multi-Source GIS Data: An Application for Landslide Risk. Remote Sensing, 14(17), 4279. https://doi.org/10.3390/rs14174279

[6] Qazvini, A. T., & Carrión, D. (2023). A spatiotemporal drought analysis application implemented in the Google Earth engine and applied to Iran as a case study. Remote Sensing, 15(9), 2218. https://doi.org/10.3390/rs15092218

[7] Van Rooijen, N., De Keersmaecker, W., Ozinga, W. A., Coppin, P., Hennekens, S., Schaminée, J., Somers, B., & Honnay, O. (2015). Plant species diversity mediates ecosystem stability of natural dune grasslands in response to drought. Ecosystems, 18(8), 1383–1394. https://doi.org/10.1007/s10021-015-9905-6

[8] Oehri, J., Schmid, B., Schaepman-Strub, G., & Niklaus, P. A. (2017). Biodiversity promotes primary productivity and growing season lengthening at the landscape scale. Proceedings of the National Academy of Sciences of the United States of America, 114(38), 10160–10165. https://doi.org/10.1073/pnas.1703928114

[9] Wang, S., Loreau, M., Arnoldi, J. F., Fang, J., Rahman, K. A., Tao, S., & De Mazancourt, C. (2017). An invariability-area relationship sheds new light on the spatial scaling of ecological stability. Nature Communications, 8(1). https://doi.org/10.1038/ncomms15211

[10] Bastos, A., Running, S. W., Gouveia, C. M., & Trigo, R. M. (2013). The global NPP dependence on ENSO: La Niña and the extraordinary year of 2011. Journal of Geophysical Research: Biogeosciences, 118(3), 1247–1255. https://doi.org/10.1002/jgrg.20100

[11] Sentinel-2 User Handbook. (n.d.). https://sentinel.esa.int/. Retrieved October 3, 2023, from https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook

[12] Level-1C Cloud Masks - Sentinel-2 MSI Technical Guide - Sentinel Online - Sentinel Online. (n.d.). Sentinel Online. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/cloud-masks

[13] Normalized Difference Vegetation Index (NDVI). (n.d.). https://ipad.fas.usda.gov/. Retrieved October 3, 2023, from https://ipad.fas.usda.gov/cropexplorer/Definitions/spotveg.htm