SOILGUARD

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

Grant Agreement no. 101000371

D5.1 Report compiling the region-specific set of evidence chains

Work package	WP5 Prediction of soil-mediated ecosystem services delivery, value and wellbeing based on soil biodiversity
Author(s)	Briony Jones (UKCEH), David Robinson (UKCEH), Chris Feeney (UKCEH), Danial Wyn Owen (UKCEH), Lindsay Maskell (UKCEH), Justine Patton (UKCEH), Alexandra Dehnhardt (IÖW), Daniel Johnson (IÖW), Katja Schmidt (UP), Laura García Velázquez (UA), Anna Clocchiatti (UVA), Pablo Sánchez Cueto (LEITAT), Laurence Jones (UKCEH)
Contributor(s)	Cross-project contributions via workshop on 7th June 2023, Athens
Reviewer(s)	Santiago Soliveres Codina (UA), Cristina Yacoub (LEITAT)
Deliverable nature	Report
Dissemination level (Confidentiality)	Public (PU)
Estimated indicated person-months	30
Date of delivery	Contractual	30-11-2023	Actual	11-09-2024
Version	2.0			
Total number of pages	34			
Keywords	Evidence chains; soil biodiversity; ecosystem services; natures contributions to people (NCP)			

Document History

Issue Date	Version	Action
20/10/2023	V0.1	Initial version
30/11/2023	V1.0	Final Version for Submission
10/09/2024	V2.0	Revised version

Table of Contents

1   Summary

This deliverable reports on progress to select, define and conceptualise evidence chains for inclusion in the Soilguardians App. Initial work focused on reviewing and incorporating ideas from existing frameworks and building on the SWBF elements. Selection and defining the evidence chains for inclusion in the App took into account a number of criteria including a clear conceptual understanding of the full chain, data and information from other WPs and external information to inform model development, and data availability for upscaling. The deliverable explains the methodologies behind the three modelling approaches required to implement different types of evidence chain in the App. This deliverable reports progress to date, and the full description of all evidence chains, as finally implemented in the Soilguardians App, will appear in D5.2.

2   Introduction

A core focus of Soilguard is to understand and represent how contrasting management of soil resources and its component biodiversity can help move towards sustainable agricultural processes, including management of grassland and forest systems. This requires understanding the interlinkages between multiple components in the soil system. The manner in which soil organisms and processes underpin or contribute to ecosystem services (or Nature's Contributions to People – NCPs) is complex, with processes operating and interacting at many levels. A wide range of management practices and soil conditions influence NCPs. For example, changes in soil compaction will influence water infiltration capacity and therefore water storage potential (Elliott and Carlson, 2004), with increased infiltration leading to reductions in overland flow (Carroll et al. 2004). Soil moisture and temperature changes will affect carbon processing rates, which in turn influence greenhouse gas emissions (Luo and Zhou, 2006), together with carbon stocks held in the soil. Faster nutrient cycling allows nutrients to become readily available for plants and may be indicative of high soil quality (Knoepp et al., 2000) but fast decomposition is likely to be detrimental for carbon storage (Luo and Zhou, 2006). Processes mediated directly by soil organisms include feeding activity by soil macroinvertebrates (notably earthworms and Collembola) influencing rates of litter breakdown (Spurgeon et al., 2005), and therefore decomposition rates and mineralization, which in turn affect nitrogen availability and carbon stocks. The cascading influences of soil and field-scale management can also influence cultural services (non-material NCPs) at larger scales. In this way, aesthetic appreciation of a landscape is likely to be higher in areas with higher plant diversity and greater abundance and diversity of forbs (Pykälä, 2003; Jones et al. 2022). Soilguard focuses on societal benefits through the frame of NCPs, so this terminology is used in the rest of this report.

The challenge which Soilguard addresses through the evidence chains is to capture key functions and processes in a coherent conceptual framework which allows an informed analysis of the linkages between individual components in a system, and their representation in a model framework which can be simplified for inclusion in a publicly accessible online tool – the Soilguardians App. The app will be used by a range of stakeholders to evaluate the change in benefits received if you change from a conventional land management method to one following organic principles (for arable systems, or the equivalent for grassland and forests).

The way that data from within Soilguard, data from external sources, and evidence from the literature are incorporated in the modelling sequence in the Soilguardians app is summarized in Figure 1. This comprises a number of elements and stages, with key data inputs also shown in the Figure. The first two stages (A and B) share the complex and inter-related set of actions, that of linking four interacting

3

components which underpin soil functions. These are soil properties and climate; land management; soil biodiversity; soil functions and processes. These stages make up a large part of the evidence chains which focuses on cascading influences within the soil system. Subsequent stages C and D represent translation of these soil functions into NCPs, with C representing the translation to biophysical units of service flow, and D representing the valuation of those service flows in economic or non-monetary terms. Stage E allows upscaling of these relationships for implementation in the App.

The approach to implement these functions in the app uses initial relationships developed as proof of concept using external data (before Soilguard-specific data is available), followed by fine-tuning of these relationships using a combination of European data and Soilguard collected data. These generalizable models can then be used to predict function, with simple input from a user of basic soil property data for their location, and climate data, both of which can be auto-extracted from existing mapped European datasets based on a provided location. This gives the app full flexibility to derive regionally or locally-specific outcomes for any location in Europe, but underpinned by Soilguard-derived more complex models and functions. These functions also allow prediction of outcomes according to management practice, e.g. a change from conventional to organic systems.

This report covers the following content: Section 3 summarises the process of conceptualizing the evidence chains; Section 4 describes prioritization of which evidence chains are key to Soilguard and should be incorporated in the App; Section 5 outlines the process of deriving relationships from existing data, adapting that to the European context, and making the links to NCPs, with section 5.1 describing Stage A, section 5.2 describing Stage B, and section 5.3 and section 5.4 describing Stage C. Section 5.5 outlines ongoing work linking to the final valuation being undertaken by WP4. Figure 1 illustrates key project data and information from each WP and how it fits into the App development.
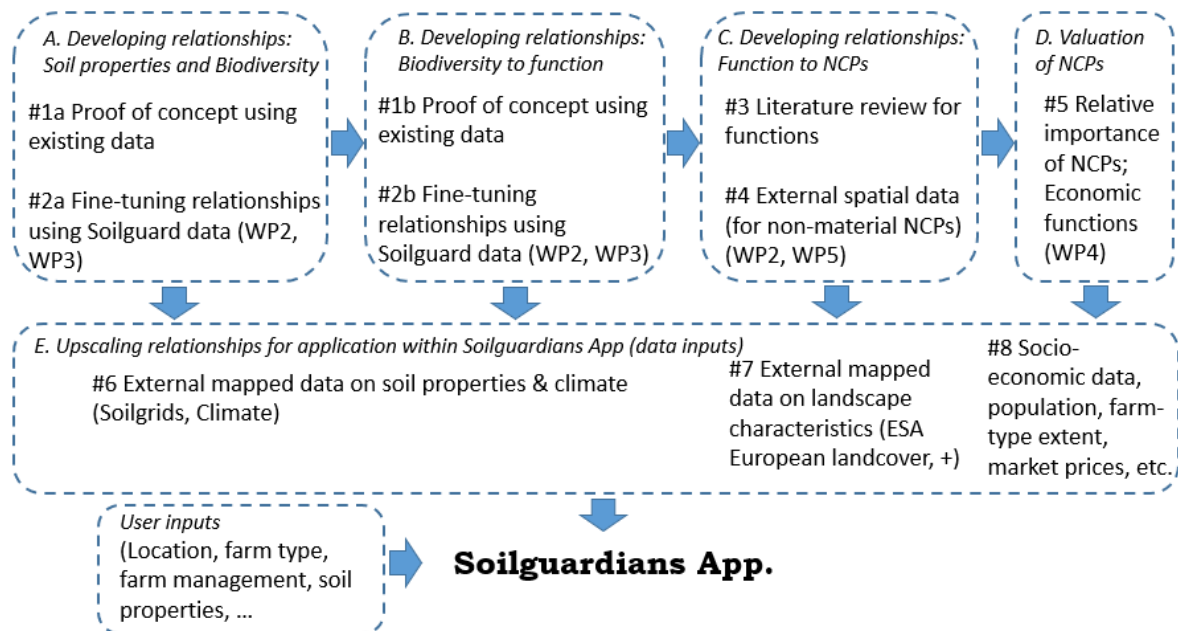

Figure 1. Elements required to implement evidence chain relationships in the Soilguardians App. A. Developing relationships linking soil properties, climate and biodiversity, B. relationships linking biodiversity and function (which may also include information on soil properties), C. Relationships linking function to NCPs, D. Valuation of NCPs. All of these require external

data for upscaling within the App to allow predictions derived from the internal models based on user input on location characteristics.

3   Conceptualising evidence chains in Soilguard

The approach taken in Soilguard builds on earlier attempts to consistently map out linkages of components of the soil system from biodiversity through to function (Hayes et al. 2018; Creamer et al. 2022). The evidence chains linking soil functions through to NCPs in large part remain a knowledge gap, which SOILGUARD is addressing. While linkages between organisms and some soil functions are becoming better documented, the linkages through to higher-level soil functions such as carbon storage are only recently being made more explicit, and the linkages to final ecosystem services (and NCPs) are not yet clear.

After reviewing a suite of key papers and resources covering this topic (e.g. Creamer et al. 2022; Drobnik et al. 2018) it became clear that despite recent advances, these efforts are only partial (i.e. they do not cover the range of organisms and functions targeted in SOILGUARD), and are often internally inconsistent (mixing up functions and services). The most consistent study in soils has set up a set of inter-linked evidence chains to demonstrate the impact of an individual stressor, in this case the trace metals of copper and mercury, and their impacts on functions and services delivered by soil organisms (Hayes et al. 2018). That study explores impacts on a range of individual organisms from bacteria to fungi to earthworms, and sets out the linkages between abundance of organisms and a series of intermediate and higher level functions and processes, ultimately leading to changes in final ecosystem services (*sensu.* Landers & Nahlik, 2013). The basic principle of an evidence chain is illustrated in Figure 2 below, showing initial influence of copper toxicity and subsequent secondary effects culminating in impact on an ecosystem service in this case. In Hayes et al. (2018), linkages between nodes were based on literature review, with key papers listed in the study to back up each relationship.
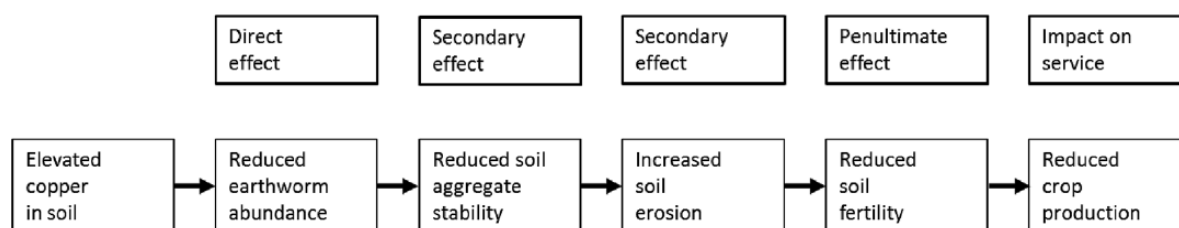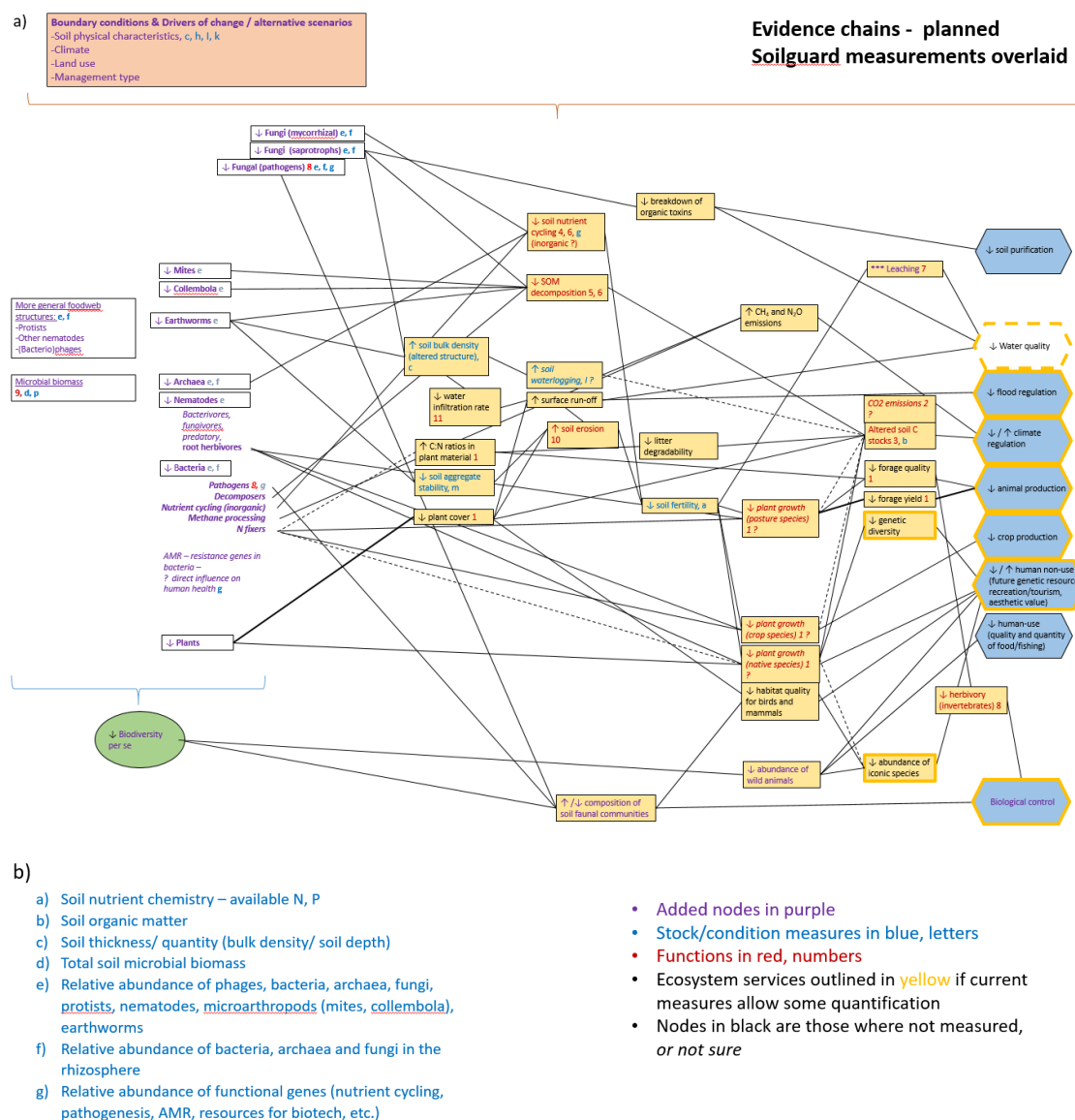

Figure 2. Components of an evidence chain illustrating direct effects, intermediate secondary effects and final impact on an ecosystem service (or NCP). After Hayes et al. (2018).

a)

Boundary conditions & Drivers of change / alternative scenarios
-Soil physical characteristics, c, h, i, k
-Climate
-Land use
-Management type

Evidence chains - planned
Soilguard measurements overlaid

b)

- a) Soil nutrient chemistry – available N, P
- b) Soil organic matter
- c) Soil thickness/ quantity (bulk density/ soil depth)
- d) Total soil microbial biomass
- e) Relative abundance of phages, bacteria, archaea, fungi, protists, nematodes, microarthropods (mites, collembola), earthworms
- f) Relative abundance of bacteria, archaea and fungi in the rhizosphere
- g) Relative abundance of functional genes (nutrient cycling, pathogenesis, AMR, resources for biotech, etc.)

- Added nodes in purple
- Stock/condition measures in blue, letters
- Functions in red, numbers
- Ecosystem services outlined in yellow if current measures allow some quantification
- Nodes in black are those where not measured, *or not sure*

1. Biomass production
2. Soil potential respiration
3. C sequestration
4. Potential N mineralization
5. Litter decomposition
6. Soil enzymatic activities (C, N, P cycling)
7. Leaching of nutrients (NO_3/PO_3)
8. Leaf (insect and fungal pathogen) damages
9. Reduction in efficiency of antibiotics
10. Soil erosion/Soil loss
11. Infiltration

NCP	Socio-economic valuation approach	Beneficiary
Food production	Market price	Farmer
Provision of timber	Market price	Farmer
Biological control	Cost-based approach (avoidance costs)	Farmer
Soil erosion prevention	Market price	Society
Nutrient cycling	Cost-based approach (replacement cost approach)	Society
Climate regulation	Cost-based approach (social costs of carbon)	Society
Flood regulation	Cost-based approach	Society
Water quality regulation	Cost-based approach (replacement cost method)	Society
Biodiversity conservation	Stated preference (monetary/ non-monetary)	Society
Aesthetic value	Stated preference (monetary/ non-monetary)	Society, Tourists

Figure 3. Modified network diagram based on Hayes et al. (2018) showing a) potential Soilguard evidence chains and (b) description of codes (please zoom in for greater detail). Black lines between nodes represent connections with thicker lines representing a greater evidence base, thin lines a lower evidence base, and dotted lines representing mixed or tentative evidence. Up and down arrows within nodes show the influence of processes on the node state, following the convention of Hayes et al. (2018) where e.g. a decrease (down arrow) in plant cover leads to an increase (up arrow) in soil erosion. Letters (and text) in blue are measures of stocks initially envisaged in Soilguard. Letters (and text) in red are measures or proxies of function initially envisaged in Soilguard. NCPs outlined in yellow can in principle be quantified in Soilguard. Nodes in purple are Soilguard specific. Text in black (excluding NCPs) are processes where quantification is uncertain.

The first stage of conceptualizing evidence chains for the purposes of Soilguard was to strip down the network from Hayes et al. (2018) to focus on linkages relevant to the project, and to add in new linkages relevant to a wider focus beyond that of toxicity impacts of trace metals. Figure 3 shows the revised network diagram which includes key input, intermediate and end-point nodes, with end-points revised to reflect the agreed focus on NCPs through WP4 (see Deliverable D4.1), rather than earlier work which used ecosystem services.

In the diagram, boundary conditions of soil type and climate, and drivers of change (See Harrison et al. 2023, D1.3) influence all elements of the system. For our purposes, the main role they play is to set the biogeographic constraints which are the setting for soil processes, i.e. basic soil properties such as soil type (which influences pH, organic matter content), climate and the main landcover/land management undertaken by a land owner at that location. Biodiversity in the chains is represented by individual groups of organisms, as well as a higher-level emergent property of biodiversity which captures elements of food-web complexity. Aspects where Soilguard can potentially provide direct or indirect measures of processes are shown in the diagram. This was used as the basis for selecting specific routes through the diagram to represent the evidence chains on which to focus for use in the App. An individual chain can then be specified on the basis of this conceptual diagram, see example shown in Figure 4. This example is hypothetical, but shows how underlying soil properties, and wider constraints such as climate, together with land management can affect both the soil organisms themselves, as well as the cascading functions.
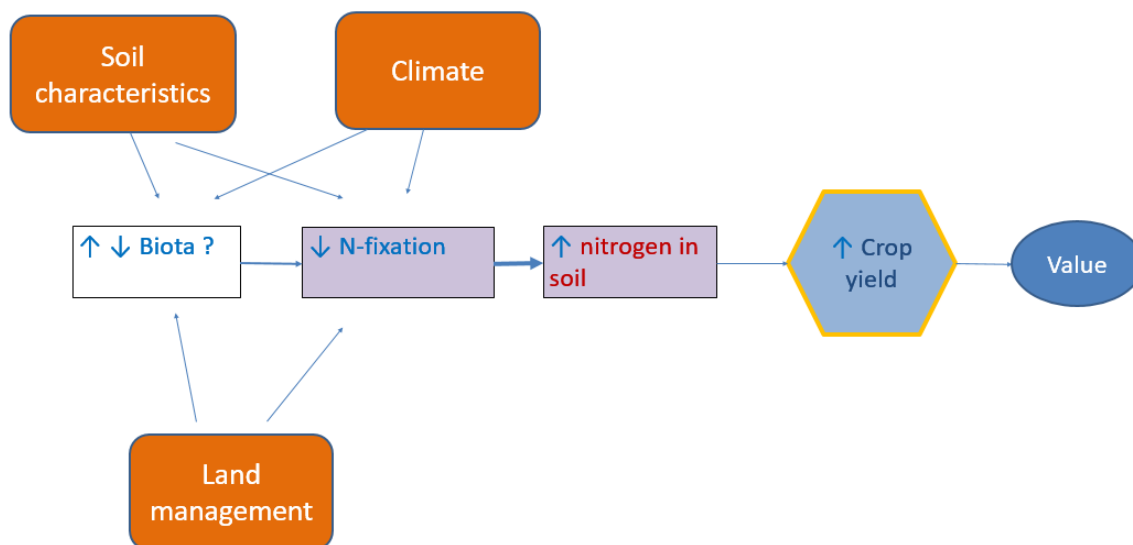

Figure 4. Illustrative example chain showing influence of background variables (soil, climate, land management) and linked nodes through to economic (or other) value. Soil characteristics include pH, organic matter content which influence e.g. microbial community composition. Climate variables include a range of potential factors such as annual rainfall, summer maximum or winter minimum temperature. Land management types include organic vs conventional as the key focus of Soilguard.

4   Prioritising evidence chains in Soilguard

The selection of evidence chains to model in Soilguard takes into account the following criteria:

- A clear conceptual link from soil properties or soil organisms through to NCPs
- Available data, response functions, or suitable proxies to allow full calculation of the chain.
- Available geospatial data which can be used to upscale calculations for the Europe domain
- Public perceptions of the importance of NCPs from WP4 surveys (reported in D4.2)
- Feedback from project steering group members

Based on these criteria, a set of six evidence chains have been prioritised for calculation in the App, across the whole European domain. The App is designed to work anywhere in Europe. It could in principle be adapted to work globally, but would require substantial rework and recalibration with different input datasets, and would only be relevant for a subset of global biomes. The first version of the App will prioritise arable systems, with other systems (grassland, forestry) planned to be added subsequently. These are summarised in Table 1, which shows the data types and methods used to calculate the chains. The conceptualised chains are shown in Figure 5 which includes the valuation method most appropriate to use, based on cross-WP discussions captured in Table A1 in Appendix 1.

The modelling approaches use three broad approaches: i) modelling of biodiversity and functional gene abundance, linked to NCPs via grouped functions described in a more detailed ontology table built from external databases (Nitrogen/Phosphorus cycling), ii) a change factor which is derived from Soilguard WP2 and other data applied to a base value derived from publicly available datasets (Crop yield, Soil carbon stock), iii) a change factor derived from spatial processing of data at a slightly larger domain scale than individual field systems (Surface water runoff/Infiltration, Soil water storage, Landscape aesthetics).

Table 1. Evidence chains linked to NCPs (right hand column), describing modelling approach, input data and intermediate linkages, indicators and functions they encapsulate. For sources of data used to implement these, refer to Figure 1 and Table A1 in Appendix 1.

Modelling approach	Input data	Intermediate linkages	Intermediate indicators	Higher level function(s)	NCP
Change factor	Location & predicted yield for that location (encompassing soil and climate characteristics)	Implicit within change model	n/a	Crop yield	**Food and feed**
Functional gene prediction	Soil pH, soil carbon or organic matter content, climate	Microbial community composition	Functional genes linked to nitrogen and phosphorus	Nitrogen cycling Phosphorus cycling	**Soil formation and protection (Nutrient cycling)**
Change factor	Soil pH, soil C or organic matter content, climate, bulk density	Implicit within change model	n/a	Soil Carbon stock	**Climate regulation**

| Change factor, linked to Hydrus 1-D modelling, and spatial analysis	Soil organic matter content, other soil properties, climate	Soil porosity	Hydraulic conductivity	Surface water runoff/Infiltration	**Regulation of freshwater quantity (Flood regulation)**
Change factor, linked to Hydrus 1-D modelling, and spatial analysis	Soil organic matter content, other soil properties, climate	Soil porosity	Water storage capacity	Soil water storage	**Regulation of freshwater quantity (Soil water)**
Change factor, linked to landscape pattern analysis	Characteristics of landscapes under different management types	Trees, hedgerows, water, field size	Abundance and composition of positive and negative landscape elements	Landscape character	**Physical and psychological experience (Aesthetic landscapes)**

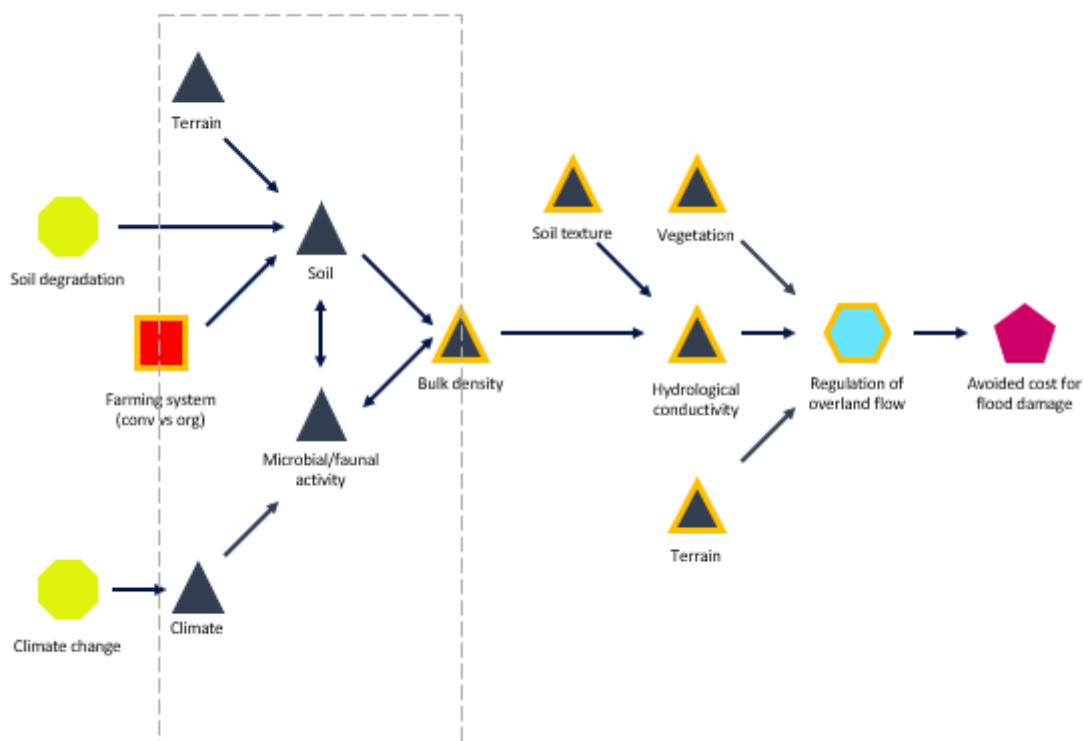## a) YIELD
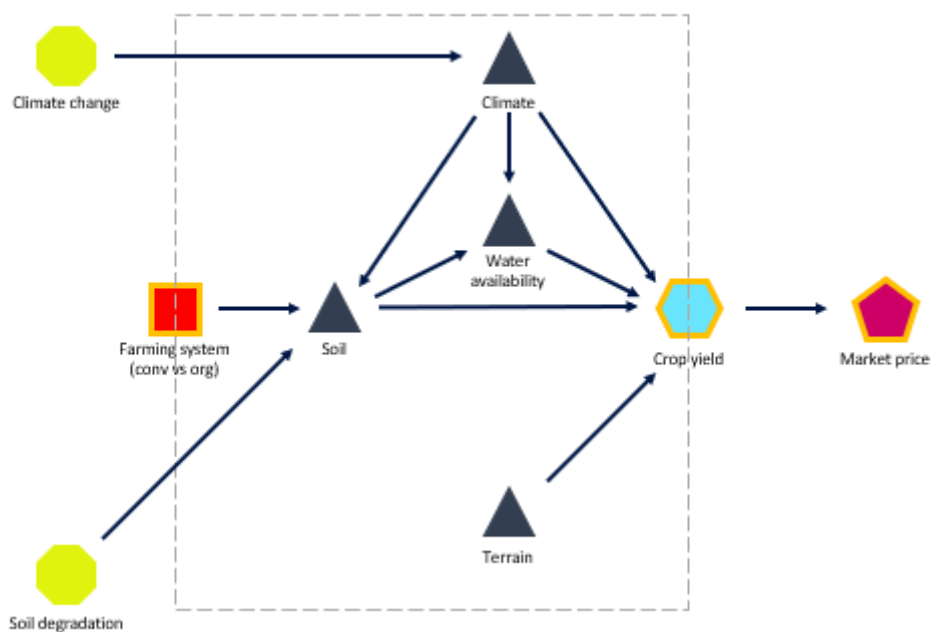
b) NUTRIENT REGULATION
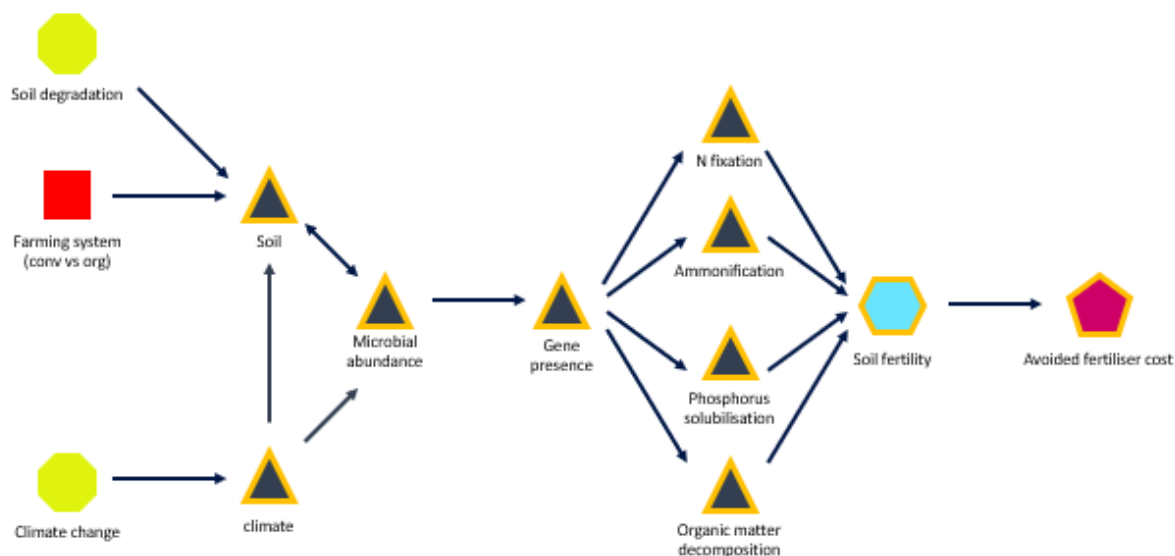
c) CARBON
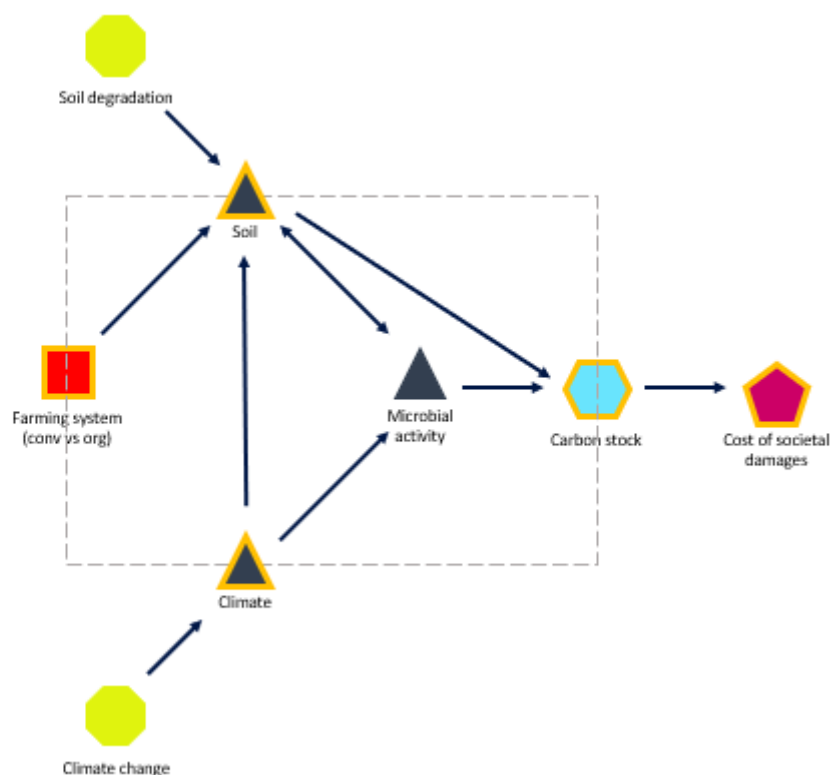
d) WATER STORAGE CAPACITY
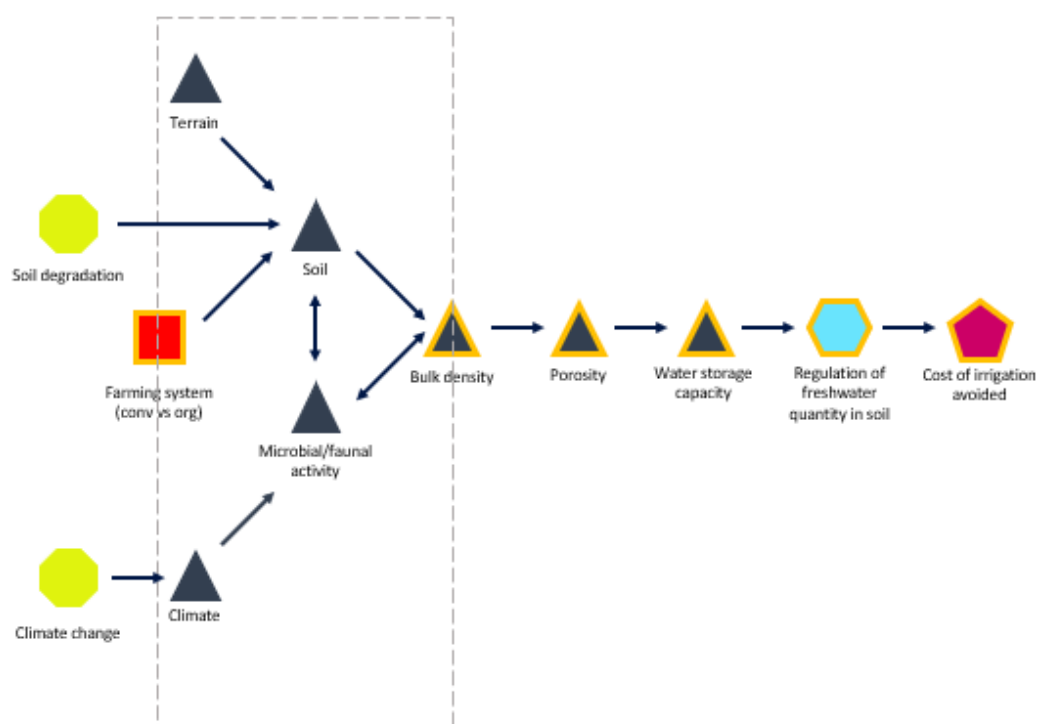
e) HYDROLOGICAL CONDUCTIVITY/INFILTRATION

f) LANDSCAPE AESTHETICS
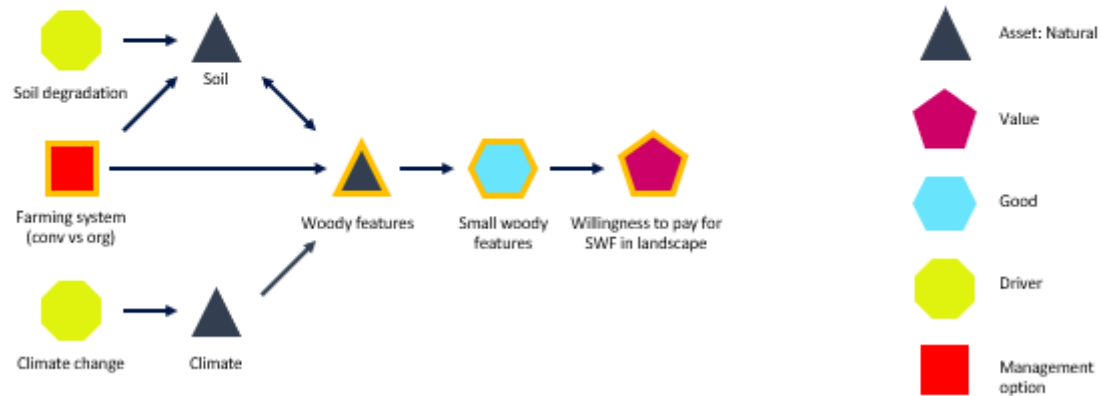
Figure 5. The six evidence chains prioritised for inclusion in the Soilguardians App. The type of node (asset, value, good, driver, management option) is indicated by shape and colour (see legend adjacent to f). Nodes between which the linkages are modelled in Soilguard are outlined in yellow. The grey dashed box outlines all linkages and nodes that are not directly modelled in Soilguard, but are implicit/already incorporated in the data used for the models. The nodes intersecting the grey box are nodes between which the linkages are modelled in Soilguard.

5   Implementing the modelling of evidence chains in Soilguard

Together, the evidence chains and the conceptual framework provide the overall construct of how we understand linked biodiversity and biogeochemical processes operating at a range of scales from molecular to field to landscape, and which ultimately influence NCPs. Since some quantification of the evidence chains is being incorporated in the Soilguardians App, this requires practical implementation of these concepts into an approach that supports calculations and representation across a wide spatial domain.

Incorporation within the App of Soilguard findings and knowledge, is planned to take a number of forms, which includes the ability to download relevant project-level data, visualisation of project-level findings and a wider prediction of outcomes for NCPs. This deliverable report focuses on the latter, i.e. to draw on ways of upscaling and moving beyond specific outcomes for sampled locations in Soilguard in order to produce more widely applicable calculation of outcomes across the European domain.

In order to be robust to the requirements of upscaling, each evidence chain needs an equation or model which can make use of publicly available and geospatial datasets at European scale. In the next sections we describe the modelling approach which underpins these calculations building on molecular biodiversity data. In a subsequent section, we outline the approach for calculating NCPs which rely on other data inputs.

The molecular modelling involves three steps (covering components of stages A, B and C in Figure 1

).
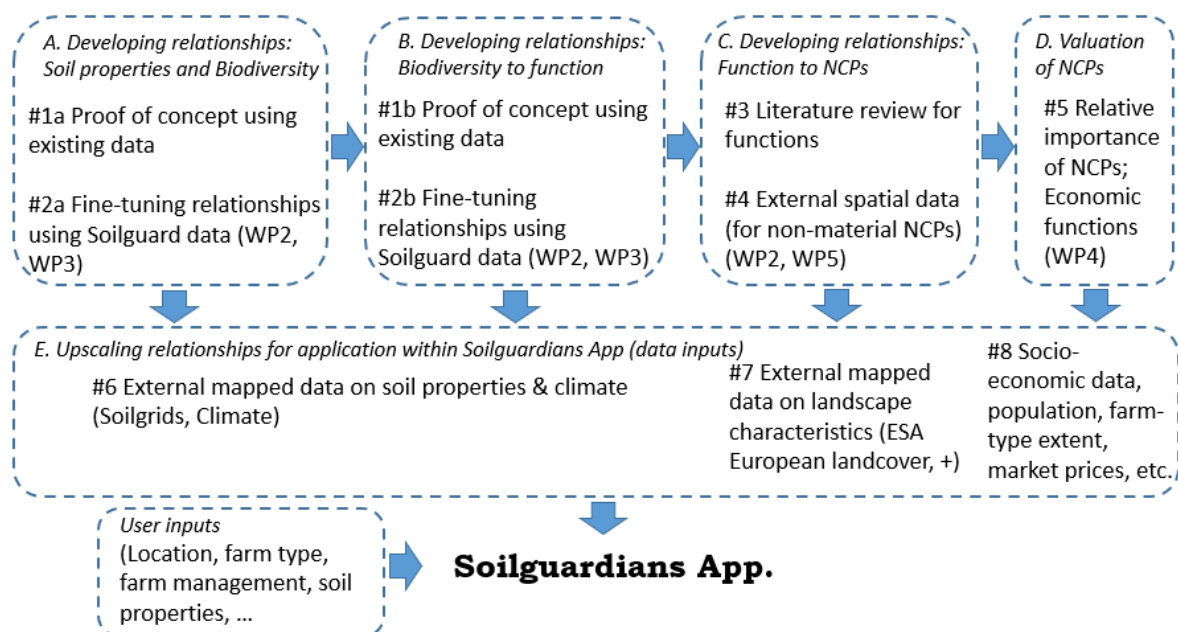
- Modelling the influence of soil and climatic properties on biodiversity composition of microbial communities.
- Modelling the links between molecular biodiversity and functional genes.
- Translating functional genes into higher-level soil functions relevant to NCPs.

13

5.1 Modelling taxonomic response to environmental change

Generalized Additive Models (GAMs)[1] (Wood, 2017) were used to establish relationships between genetic measures of organism relative abundance, and biophysical parameters linked to soil characteristics.

5.1.1 Preliminary taxonomic modelling approach

In Preliminary analyses, GAMs were modelled with an external dataset to establish proof of concept, using the Countryside Survey (CS) dataset, a UK distributed soil survey consisting of 1152 soil samples incorporating a diverse range of habitats from agricultural to semi-natural, including woodland. A GAM modelling approach was chosen due to their easy interpretability within an ecological context, with GAM visualisation methods enabling understanding of specific niche preferences. GAMs were generated using 16S marker gene data[2] on the OTU level[3]. Prior to modelling, samples with less than 5000 DNA sequencing reads & OTUs that were present in less than 30 samples were discarded from analyses, resulting in a filtered dataset of 993 samples and 7046 OTUs

The R mgcv package was then used to fit GAM distributions of the top 1000 most dominant taxa (of the remaining 7046 OTUs). Initially, smooth functions of pH and LOI and an interaction term were used as predictors of rarefied[4] OTU abundance. These models use LOI as an input, which can be calculated from soil C values by applying the conversion factor of 1.724. A negative binomial family was used and restricted maximum likelihood REML method was selected for parameter smoothing estimation. For initial validation of the approach CS data was split into a training and test subsets of 40% and 60% respectively. Once the GAMs were assessed to be a viable method the models were rerun on the complete CS dataset and further validated with independent data. Ongoing refinements of the approach include incorporating climate variables as additional predictors, and translating the modelling approach to make use of Soilguard and LUCAS data. These will directly use soil %C data as input in place of LOI.

5.1.2 Validation of GAM approach

Model validation was conducted using UGRASS, an independent land use intensification soil survey carried out in the UK, funded by the Soil Security Programme. A paired sampling method was undertaken whereby adjacent low land use intensity sites (pristine grasslands) and high land use intensity sites (improved grassland/ arable sites) were sampled, thereby controlling for underlying soil and climatic gradients to isolate factors attributable to management intensity. UGRASS data was pre-

---

[1] Generalized Additive Models (GAMs) are a statistical modelling technique that allow for non-linear relationships between response variables and predictors.

[2] 16S marker gene analyses– molecular methodology whereby small subunit rRNA is used to determine bacterial community composition.

[3] OTU (Operational Taxonomic Unit)- groupings of closely related taxa determined by taxonomic marker gene sequence similarity. Within this work OTUs were generated using a sequence similarity threshold >=97%.

[4] Rarefaction- Normalisation method used for DNA sequencing data, specifically data is resampled to standardise number of DNA sequencing reads across samples.

processed by discarding samples with <=5000 sequencing reads, and then normalising read numbers by rarefying to the lowest remaining read number across samples. To identify CS modelled taxa within the UGRASS dataset, UGRASS OTU representative sequences were aligned against Countryside Survey representative sequences using blastn[5] with an evalue[6] of 0.001. Predictions were then made for taxa with a >=97% identity[7] hit against a CS OTU with an associated GAM model. Performance of models within the independent data was assessed both on the community and individual taxonomic level. On the community level sample similarity was calculated within both observed and predicted datasets using the Vegan R function VegDist. Sample distances within the observed and predicted data were then correlated and demonstrated a mantel r test statistic of 0.785 (Figure 6a). This strong correlation suggests pH and LOI are able to predict broad bacterial community metrics.

---

[5] Blastn - Bioinformatic tool comparing sequence similarity of query DNA sequences to a reference database.

[6] Blast evalue – expected number of sequence alignments of the same or greater quality that could occur by chance given the size of the reference database.

[7] Blast identity- level of sequence similarity between a query and reference sequence- specifically the number of matching nucleotides aligned between the query and the reference.
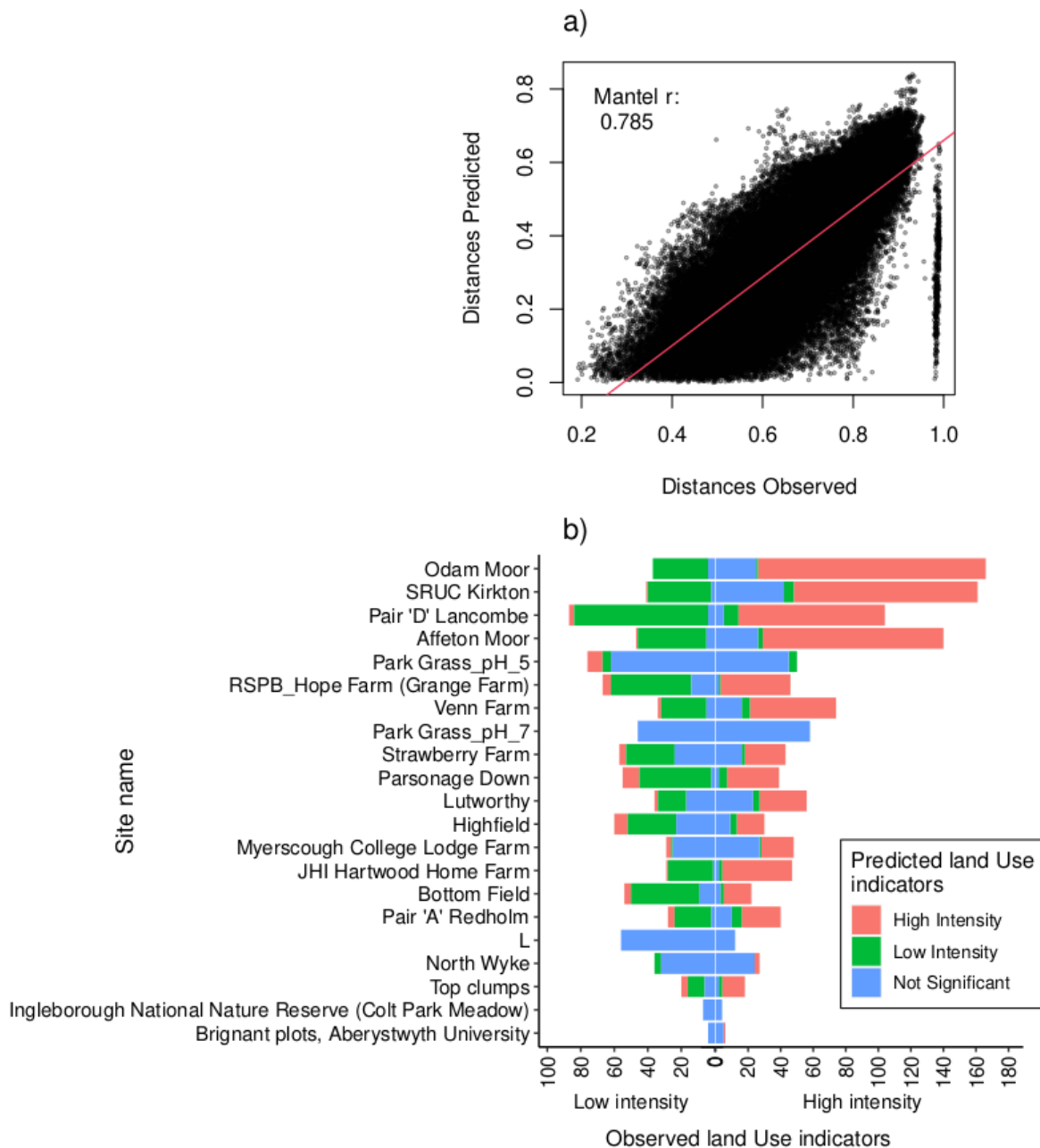
Figure 6. Validation of Generalized Additive Models (GAMs) approach for modelling bacterial data. GAMs were generated using Countryside survey bacterial marker gene data (16S). Individual bacterial taxon (OTU's-Operational Taxonomic Units) were used as response variables and pH and LOI as predictors with interaction (OTU~ s(pH) + s(LOI) + ti(pH, LOI)). GAMs were validated using an independent 16S dataset from the UK land use intensification survey UGRASS. The figure shows a) predictions on the community level, where sample dissimilarity was calculated separately for observed and predicted 16S communities using the R Vegan function VegDist. Sample dissimilarity scores for observed and predicted data were then correlated using mantel test statistic, and b) evaluation of the model's ability to recall specific taxonomic responses, where land use indicators were identified in observed and predicted data using Dufrene legendre indicator analysis. Bars depict the observed quantity of high intensity and low intensity land use indicators within each site, whilst colours represent the predicted indicator class.

To test whether the models were also able to correctly capture responses at the individual taxonomic level, Dufrene Legendre indicator analyses[8] was used to identify per site indicators of land use intensity within observed and predicted data. We observed that the modelled data successfully recalled a substantial proportion of observed high and low land use intensity indicators within most sites (Figure 6b). As expected for some specific sites such as the Park Grass experiment where liming was used to modulate soil pH and variations in organic matter were small, model performance was comparatively poorer.

5.1.3   Validating GAM approach with European data

We next sought to assess how well models built on GB data performed on European data. To do this we validated the data on the Ecofinders transect consisting of ~80 soils encompassing 11 European countries, featuring woodland, grassland and arable soils. Ecofinders data was pre-processed by removing samples with <=5000 sequencing reads, and then normalising reads by rarefying to the lowest read count across samples. We subset the Ecofinders OTUs to those shared with the modelled subset of the CS dataset (top 1000 abundant CS taxa). To identify Countryside Survey (CS) modelled taxa within Ecofinders, Ecofinders OTU representative sequences were aligned against Countryside Survey representative sequences using blastn with an evalue of 0.001. Predictions were then made for 1537 Ecofinders taxa with a >=97% identity hit against a CS OTU with an associated GAM model. The performance of models was assessed both on the community and OTU level. On the community level sample similarity was assessed within both observed and predicted Ecofinders data using the Vegan R function VegDist. Sample distances within the observed and predicted data were then correlated and demonstrated a mantel r test statistic of 0.843 (Figure 7a). To assess the prediction on the OTU level an NMDS ordination[9] object was created for observed and predicted data, NMDS1 scores for each OTU were then calculated for both datasets and correlated using spearmans correlation (Figure 7b). Generally these NMDS1 scores correlated well (spearmans r 0.821) though some outliers can be observed (figure 2b). We next sought to see if global land use indicators could be recalled within the predicted data, through conducting dufrene legendre indicator analysis on both observed and predicted data.

---

[8] Dufrene Legendre indicator analyses – Form of indicator analyses used to identify variables (here bacterial taxon) responsive to specific sample types (here land use). Dufrene Legendre takes into account both the specificity of a variable (i.e. here how specific a bacterial taxon it is to a particular land use) and fidelity (i.e. here how consistently a bacterial taxon occurs within that sample type).

[9]   Ordinations- summarises community data in low dimensional space to simplify variance in the data. Arrangement of taxa or samples relate to level of similarity.
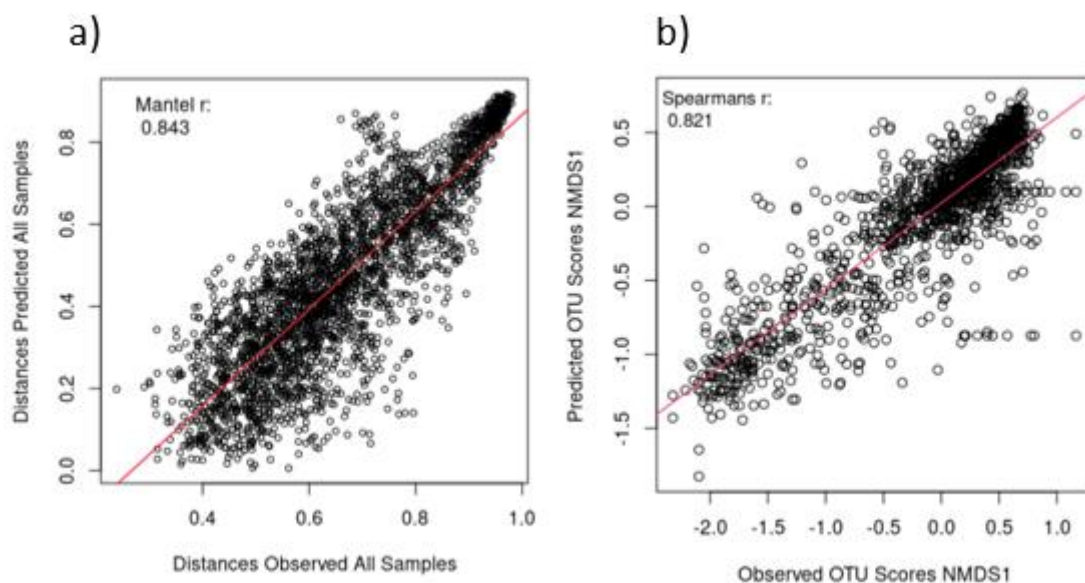
Figure 7. Validation of bacterial Generalized Additive Models (GAMs) on European level data. Here GAMs were validated using an independent European 16S dataset from the Ecofinders transect. a) To assess predictions on the community level, sample dissimilarity was calculated separately for observed and predicted 16S communities using the R Vegan function VegDist. Sample dissimilarity scores for observed and predicted data were then correlated using mantel test statistic. b) To evaluate the accuracy of individual OTU predictions, observed and predicted NMDS scores for each OTU were compared and correlated using Spearmans correlation.

As seen in Table 2, a large proportion of indicators could be successfully recalled particularly for arable (80.69%) and forestry sites (86.29%), though a smaller percentage was successfully recalled for grassland (57.25%). Together these findings suggests that the GB models perform reasonably well on European data. This model performance will be further improved with inclusion of climate variables into the GAM predictors.

Observed indicators

	Arable indicator	Forestry indicator	Grass indicator	Not Significant
Arable indicator	80.69	1.61	23.91	40.26
Forestry indicator	0.86	86.29	2.17	12.9
Grass indicator	6.01	3.23	57.25	25.09
Not Significant	12.45	8.87	16.67	21.74

Predicted indicators (row label on left side)

Table 2. Validation of Microbial GAMs using European dataset Ecofinders at the taxonomic level. Land use indicators were calculated for observed and GAM predicted European data on the OTU level. The percentage of observed indicators predicted to be each indicator class were then assessed.

5.1.4   Ongoing model development

Having built preliminary taxonomic models with UK data and validating the approach, we have now rebuilt models on SoilGuard WP2 data released to other work packages. Individual taxonomic abundancs were modelled on the ASV (Amplicon Sequence Variant) level using pH, OM, annual rainfall (WorldClim2 BIO12) and mean temperature of the warmest quarter (WorldClim2 BIO10) as predictors (Fick & Hijmans, 2017). These specific predictors were chosen in preliminary analyses based on random forest variable importance statistics.

Bacterial marker gene data was normalised through rarefication by WP2. Taxa that were present in less than 30 samples were discarded from analyses and were considered to not have a high enough sample occupancy to meaningfully model. The 100 most abundant taxa from each sample were then selected to model, to provide good representation of communities across sites and countries. A total of 3195 unique taxa were then modelled using GAMs (Wood, 2017) and the mgcv package with a negative binomial family.

5.2   Modelling relationships between organism relative abundance and gene functional potential.

5.2.1   Modelling approach

In order to link context specific changes in taxonomic communities to functional potential, a set of random forest models[10] (Breiman, 2001) were generated predicting functional genes from bacterial taxonomic composition. This first proof of concept used data from extensive surveys using data from Great Britain.

Models were generated using collocated bacterial marker gene and metagenomic[11] data from the UGRASS land use intensification survey. Metagenome read numbers were normalised by rarefying to the lowest read number across all samples. Genes that were not present in at least 30 samples were discarded. Bacterial marker gene data was pre-processed by removing samples with sequencing a read number of <5000 and then rarefying to the lowest read number across sites. To enable GAM models and random forest models to be easily chained (to contribute to operationalized evidence chains) only UGRASS OTUs with a >=97% hit to the GAM modelled CS OTUs were retained. UGRASS data was randomly split into training and test subsets using a 75%/25% split (consisting of 54 and 20 samples respectively). Random forest regression models were generated per gene using the rarefied gene abundance as the response variable and the observed 16S taxa as predictors.

5.2.2   Model validation

Random forest models were validated through generating two sets of functional gene predictions on the test subset of the UGRASS data. Prediction i) used observed UGRASS 16S data as predictors and prediction ii) used the GAM predicted 16S abundances. Performance of random forests were tested

---

[10] Random forest models- machine learning approach combining numerous decision trees with each decision tree consisting of a subset of samples and predictor variables.

[11] Metagenomics- methodology whereby DNA is directly sequenced from a sample capturing the collective functional potential of a microbial community.

on the community and gene level. On the community level the R Vegan function VegDist was used to assess sample similarity within the observed and predicted metagenome datasets. The sample distances within observed and predicted data were then correlated using the mantel test statistic.

A strong correlation (mantel r test statistic 0.707) was seen between observed metagenome sample distances and prediction i sample distances, suggesting it is possible to predict broad metagenome profiles from observed 16S communities (Figure 8a). Similarly observed sample distances also correlated well with prediction ii sample distances with a slightly weaker mantel r test statistic of 0.571 (Figure 8b). This comparatively weaker correlation is unsurprising given than prediction i predicts functional genes from observed OTUs and prediction ii predicts functional genes using GAM predicted OTUs. Nevertheless, the reasonably strong correlation demonstrated does suggest that chaining the GAM with RF models described is likely a viable approach. To assess performance on the gene level, observed and predicted gene relative abundances were correlated using Spearman's correlation (Figure 8c). Observed vs prediction i gene relative abundances demonstrated good levels of correlation (blue histogram) with a mean observed vs predicted i r^2 of 0.577 and 40.3% of genes possessing a r^2 of >0.7. Observed vs predicted ii relative abundances (red histogram) also correlated well with an average r^2 of 0.43 and 17.1% of genes possessing a r^2 of >0.7.

In order to establish if the same genes were predicted with a comparable level of accuracy between approaches, r^2s between predicted and observed genes were correlated between approaches, resulting in a Spearman's correlation of 0.798 (Figure 9). This finding that the same genes that could be accurately predicted using approach i could also be accurately predicted using approach ii, further demonstrates the viability of chaining GAM and RF models.
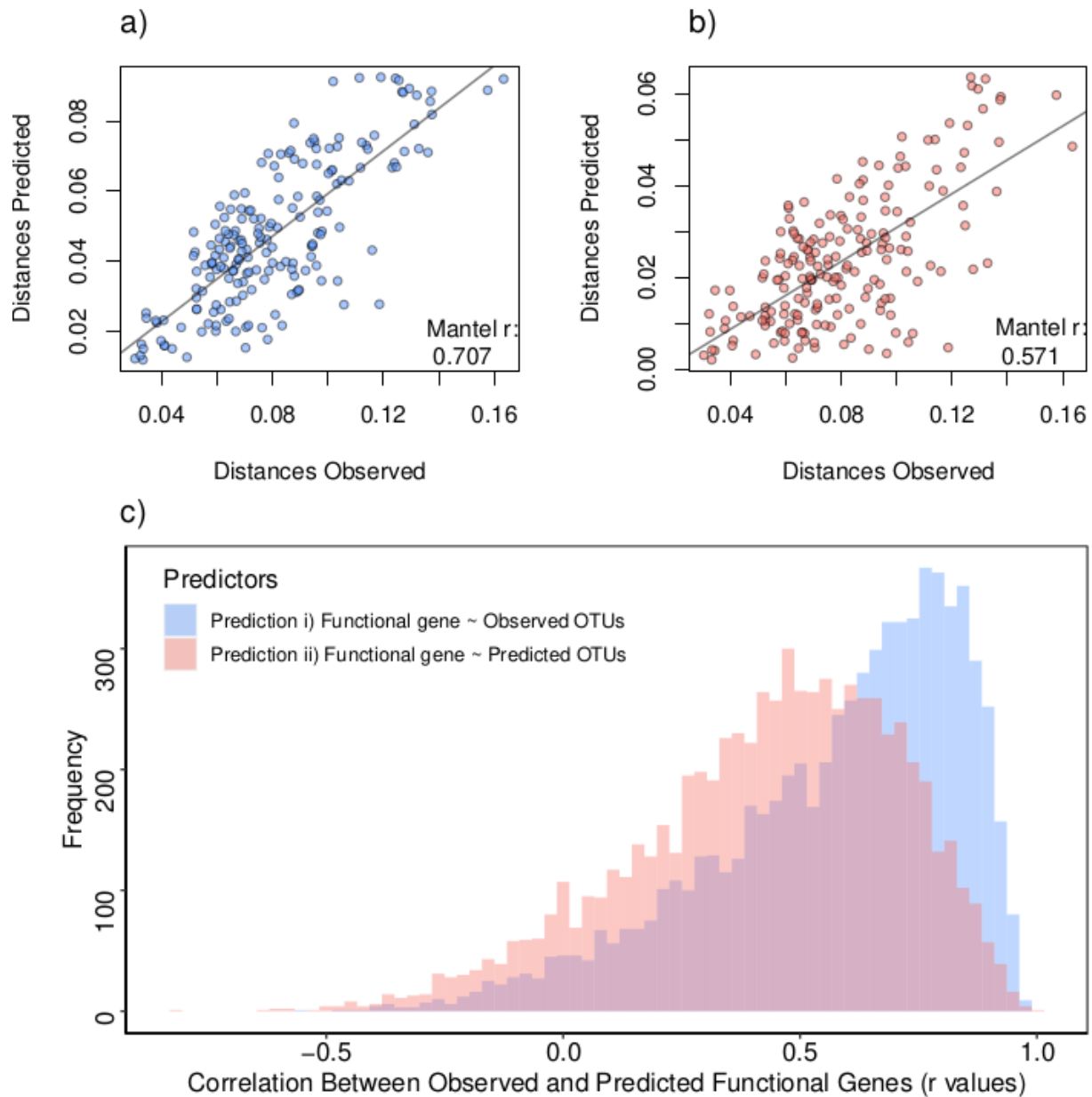
Figure 8. Validation of random forest models, predicting gene relative abundances from bacterial communities. Random forest models were generated using the UK land use intensification survey UGRASS using collocated bacterial marker gene (16S) & metagenomic data. Models were then validated on an independent subset of the UGRASS data. Two prediction approaches were used: prediction i - predicted functional genes from observed 16S data and prediction ii - predicted functional genes using GAM predicted 16S data. To assess the accuracy of the broad predicted functional profiles, sample dissimilarity indices were calculated for observed and predicted data in an independent subset of the UGRASS data and then correlated using the mantel test statistic. a) Depicts the correlation between observed and prediction I (predicted from observed OTUs) sample indices b) shows the correlation between observed and prediction ii (predicted from predicted OTUs) sample indices. To assess the accuracy of predicted responses at the gene level observed and predicted genes were correlated using spearman's correlation, the stacked histogram c) shows the distribution of gene correlations, with histogram colour representing the prediction approach used.
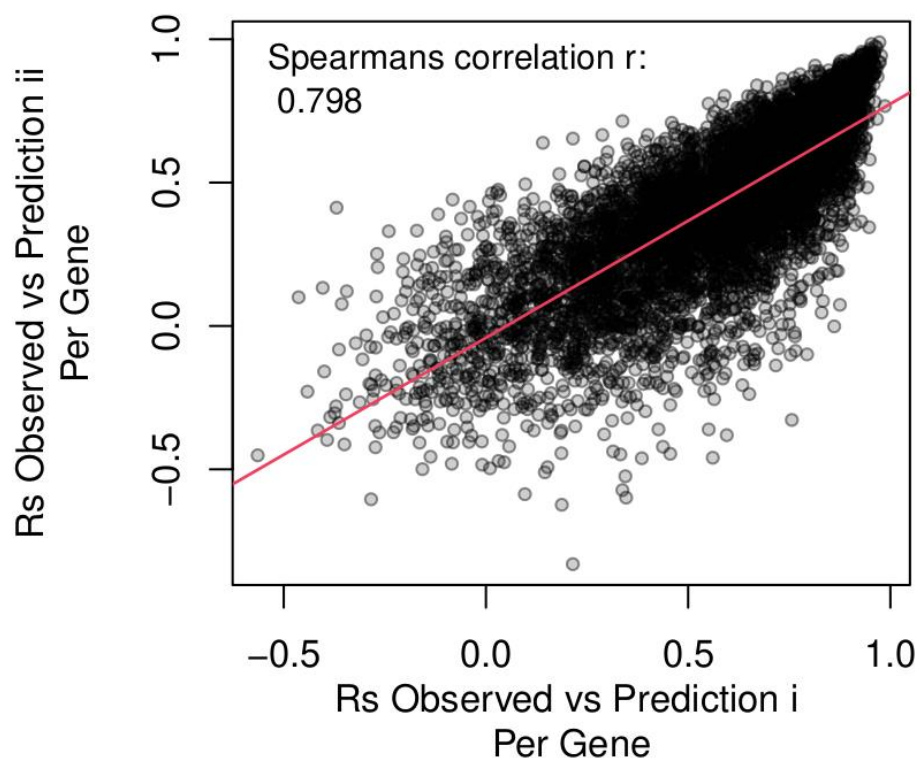
Figure 9. Consistency between metagenome prediction approaches. Metagenome data was predicted using two approaches, prediction i relative abundances were predicted from observed OTUs whilst prediction ii relative abundances were predicted using GAM predicted OTUs. For each gene, observed relative abundances were correlated separately with the two sets of predicted relative abundances, generating two lists of per gene r values for each approach, these r values were then correlated to assess if the per gene accuracy showed consistency between approaches.

The predictive approach to soil microbial communities is described in a manuscript, currently in draft. The paper demonstrates not only that pH and LOI can predict broad community metrics, but also can predict specific taxonomic responses and recall indicators of land use intensification. It demonstrates the proof of concept linking bacterial taxonomic communities to functional genes, showing that amplicon data can be used to predict functional profiles of communities at the broad level and to predict relative abundance of specific genes including correctly predicting the relationship between land use intensification and nitrate reductase genes (known consistent arable indicators).

5.3   Linking functional genes to broad function & NCPs

5.3.1   Curation of gene functional mapping table

To link normalised gene abundance to broad soil functions and or Nature's Contributions to People (NCPs) we next sought to interrogate gene ontology annotations. The SEED subsystem ontology database was used (Overbeek et al., 2005), though SEED has not specifically been curated for environmental research it contains a number of relevant functional categories such as "Nitrogen

metabolism", "Phosphorus metabolism" and "Carbohydrates". The database also possesses a hierarchical structure whereby each gene can be linked to up to four levels of subsystems, with level 1 describing the broadest functional/structural category and level 4 describing groupings in the greatest level of detail, thus enabling subsets of genes to be identified at various levels of functional specificity.

Outputs from the WP5 Value Chains Workshop (held at the SoilGuard 24M meeting in Athens), were used to establish which genes/subsystems were of relevance to SoilGuard. The WP5 workshop table has since been refined, shown in Appendix 1. This was used as the basis for links between soil properties measured in WP2/WP3 to NCPs in WP4/WP5 and was partially populated by active participants of the workshop. SEED level 1 and level 2 subsystems were manually checked for groupings relevant to each measurement and NCP listed within the WP5 workshop table (see Table A1, Appendix 1). Using this approach, a total of six initial SoilGuard-relevant functional categories were defined, including "Carbon cycling", "Nitrogen cycling", "Phosphorus cycling" and "Heat/Osmotic_Stress", with each category consisting of genes within one or more related level 1 or level 2 subsystems. A gene functional mapping table was then curated using these categories. Genes of relevance to one or more categories were designated a general relevance code of 1, whilst those not of relevance to any functional category were given a code of 0. Each gene was also designated a relevance code for each specific functional category with a value of 1 indicating relevance to that particular function.

Of the 7,061 modelled functional genes, 2,218 (31.4%) were classified as relevant to one or more of the functional categories described. Most functionally relevant genes were related to a single function and just 0.81% (18 genes) were classified as relevant to multiple functional categories. Future development of this table will involve refining linkages between genes and SoilGuard relevant functions though further interrogating gene ontology databases, with the immediate priority to consider level 3 SEED Subsystem annotations. This will allow a more refined set of mapping across of genes to function, for example allowing differentiation between components of the carbon cycle and of the nitrogen cycle, and a more nuanced interpretation of how functions link to NCPs in the subsequent phases of implementing this approach in the Soilguardians App. The high level functions are then further combined into relevant NCPs. For example, both N cycling and P cycling contribute to the NCP of nutrient cycling. These linkages will be further refined with more in-depth analysis of sub-systems mapping for individual genes in order to separate out, for example, genes which speed up carbon cycling and genes associated with processes which indicate conditions which are associated with slower rates of carbon turnover. At this stage, the heat/osmotic stress function was considered out of scope for the App.

5.4   NCPs requiring other data inputs

Two of our selected NCPs require separate datasets for calculation purposes.

For water infiltration, we base the approach on studies linking water retention in landscapes to soil physical and chemical parameters. The study by Robinson et al. (2022) provides a response function, which allows a quantified link for soil porosity as a function of soil organic matter content, and the function is robust against a wide range of land cover types, from arable to woodland (Figure 10). This allows upscaled predictions through use of European soil maps and WP2 data which characterises bulk density and other soil properties into different classes, summarising potential for water storage within

23

soil (see Appendix 1). This last step is necessary to allow us to differentiate conventional and organic in the predictions of the App.
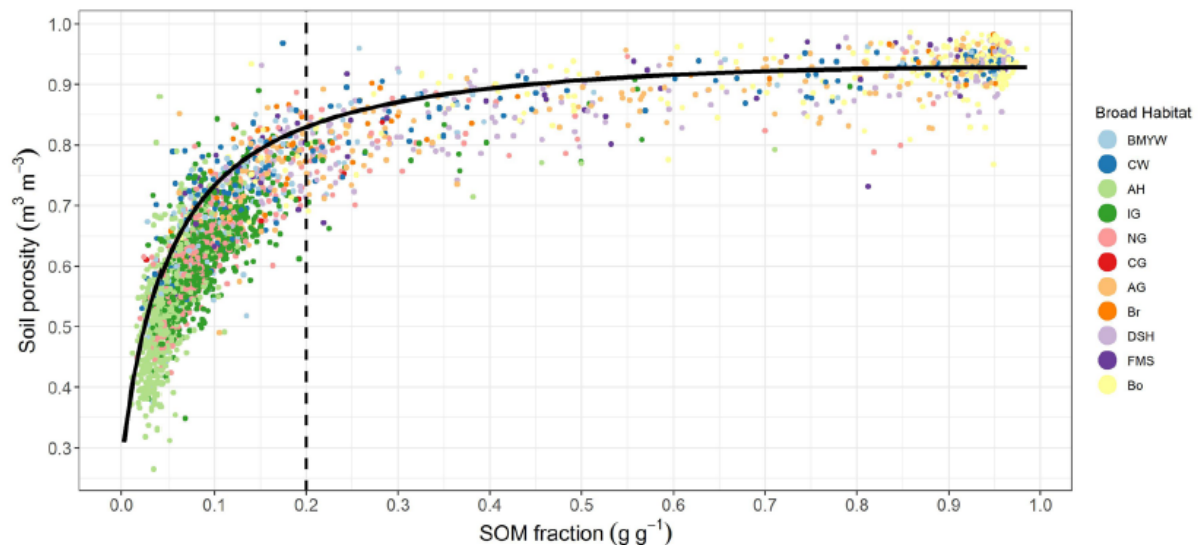

Figure 10. Relationship between soil organic matter and soil porosity. Relationship from Robinson et al. (2022). Broad habitat codes are: BMYM – Broad leaf, Mixed and Yew woodland; CW – Coniferous Woodland; AH – Arable and Horticulture; IG – Improved Grassland; NG – Neutral Grassland; CG – Calcareous Grassland; AG – Acid Grassland; Br – Bracken; DSH – Dwarf Shrub Heath; FMS – Fen, Marsh and Swamp; Bo – Bog.

For landscape scale aesthetics, we need to think at a landscape scale, which brings in the need for additional datasets, and the requirement to consider the linkages between conventional and more sustainable agricultural practices in a different way, noting that the majority of the data we hold in Soilguard is at field-scale.

Modelling landscape character includes a number of components which may be influenced by agricultural methods at a field and landscape scale. For example, certain features can be considered as either positive or negative in a landscape context. Natural elements such as trees and water tend to lead to positive appreciation of character (Tveit et al. 2006) while negative features tend to include some, but not all, human-made features. In particular, industrial and built features are predominantly seen as negative, although once industrial features acquire historic or cultural standing, then they can become positive features (Koblet & Purves 2020). Openness in landscapes features strongly in the literature as being a positive element in landscapes (Schirpke et al. 2013). Openness can be conceptualised as having a lack of view-hindering features and as the presence of views across a wider landscape with other features – such as views of mountain peaks or distant horizons (Hedblom et al. 2019), but can also be culturally determined. People growing up in very open landscapes with little habitat diversity, large fields and wide horizons view these elements positively, while the more general perspective is that people intrinsically value more diverse landscapes higher than less diverse landscapes. Views relate to visual scale in the landscape (Tveit et al. 2006) can be referenced back to both evolutionary and aesthetic theories about human responses to landscape (Kaplan & Kaplan 1982, Bell 1999).

24

Within Soilguard, we will focus modelling approaches on landscape elements most sensitive to the land management practices considered in the project, combined with availability of suitable data. The approach will make use of landscape character mapping methods developed by Fletcher et al. (2022), and incorporate indicators covering linear features in the landscape, field size, degree of woodland cover and water in the landscape, building on foundational work undertaken in WP2 on the importance of landscape elements in different size buffer zones around WP2 sampling locations.

5.5   Applying economic and other values to NCP outcomes.

Appropriate ways of valuing NCPs differ both according to the type of NCP, and according to the valuation evidence that is collected within Soilguard, or which can be derived from external sources. The types of valuation include market values for NCPs linked to products such as crop or timber yield, some market values for climate regulation (valuation of carbon stocks or sequestration), non-market values for some regulating services (such as replacement costs for fertiliser use) and ratings (preference values) where economic valuation is not possible or not appropriate (applicable both to regulating material NCPs and non-material NCPs like aesthetic landscapes. Market values and some other cost-based measures can be applied to any location in Europe in the app. Ratings developed in the project may only be applicable to specific regions where data is being collected during Soilguard. Valuation options are summarised in Table A1 in Appendix 1.

6   Conclusions

This report outlines the process of defining and selecting the evidence chains to be represented in the Soilguardians App, following criteria of clear conceptual linkages between soil properties and final NCPs, data available within the project or externally to compute the chains, and data available which allow upscaling to wider European predictions within the App. Through this process, four chains were selected, spanning a range of non-material NCPs. While it was possible to quantify the material NCP of crop-production for some components of WP2 sampling sites, the methodology currently does not allow upscaling of e.g. crop yield across the Europe domain, so this NCP is not included in the wider Europe-focused App functionality. However, in the remainder of the project we will explore options to incorporate a proxy for this element, for example using basic relationships differentiating yield in conventional vs organic systems derived from Soilguard data, and applying these to globally available maps of yield for a specific crop type. This would be possible for a limited number of crops to indicate relative changes in yield, but would not be possible to implement a fully upscaled version applicable across all Europe for multiple crops. Separate information on differences in crop (or grassland/timber) yield when changing between management systems will be available for reference from other Soilguard data stored in the app.

The report illustrates the calculation and modelling processes which underlie most linkages in the evidence chains. These range from models based on microbial community composition derived from molecular data, and models linking community composition to relative abundance of functional genes,

to response functions linking soil chemistry to soil porosity, and finally landscape-scale metrics which can be calculated at much larger scales and relate to non-material NCP of aesthetic appreciation.

Remaining steps include: As soon as full collated Soilguard data is available, this will be used to fine-tune relationships currently derived from external UK and European datasets. Further differentiation of gene functions by in-depth breakdown of mapped sub-systems allocated to individual gene functions will be used to derive more nuanced relationships linking genes to the high level functions which can map to NCPs. Further work is also required to scale these relationships to quantities which can be expressed per unit area or per unit quantity of a process (e.g. change in carbon stock per ha), or rate over time, which allow more direct valuation for those NCPs using market costs or other costs requiring unit values. This work will also fine tune the relationships with those NCPs to be calculated as ratings, for example to establish a relative scale against which to evaluate change. This will build on the separate benchmarking calculations, which put a given location into its wider European context, based on the soil properties and climate conditions at that location compared with the full European distribution.

References

Bell, S. (1999) Landscape: Pattern, Perception and Process (London: E&FN Spon).

Breiman L. (2001). Random Forests. Machine Learning 2001; 45: 5-32.

Carroll, Z.L., Bird, S.B., Emmett, B.A., Reynolds, B., Sinclair, F.L. (2004). Can tree shelterbelts on agricultural land reduce flood risk? Soil Use Manage. 20 (3), 357-359.

Creamer, R.E., Barel, J.M., Bongiorno, G. and Zwetsloot, M. (2022). The life of soils: Integrating the who and how of multifunctionality. Soil Biology and Biochemistry, p.108561.

Drobnik, T., Greiner, L., Keller, A. and Grêt-Regamey, A. (2018). Soil quality indicators–From soil functions to ecosystem services. Ecological indicators, 94, pp.151-169.

Elliott, A.H., Carlson, W.T. (2004). Effects of sheep grazing episodes on sediment and nutrient loss in overland flow. Aust. J. Soil Res. 42 (2), 213-220.

Fick SE, Hijmans RJ. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 2017; 37: 4302-4315.

Fletcher, D., Norton, L., Jones, L. (2022). 'Changes in landscape and waterscape character' Development of G1 indicator for Defra 25YEP. Change mapping report. Report to Natural England, March 2022.

Harrison, P.A., Maskell, L., Robinson, D., Jones, L., Griffiths, R., Lladó, S., Olivera, S., Hartmann, M., Dehnhardt, A., Yacoub, C., Muñoz, M. (2023). Soil Biodiversity and Wellbeing Framework. Deliverable D1.3 for EU H2020 Soilguard project.

Hayes, F., Spurgeon, D.J., Lofts, S. and Jones, L. (2018). Evidence-based logic chains demonstrate multiple impacts of trace metals on ecosystem services. Journal of environmental management, 223, pp.150-164.

Hedblom, M., Hedenås, H., Blicharska, M., Adler, S., Knez, I., Mikusiński, G., Svensson, J., Sandström, S., Sandström, P. and Wardle, D.A., 2020. Landscape perception: linking physical monitoring data to perceived landscape properties. Landscape research, 45(2), pp.179-192.

Hollis D, McCarthy M, Kendon M, Legg T, Simpson I. (2019). HadUK-Grid—A new UK dataset of gridded climate observations. Geoscience Data Journal 2019; 6: 151-159.

Jones, L., Boeri, M., Christie, M., Durance, I., Evans, K.L., Fletcher, D., Harrison, L., Jorgensen, A., Masante, D., McGinlay, J. and Paterson, D.M. (2022). Can we model cultural ecosystem services, and are we measuring the right things?. People and Nature, 4(1), pp.166-179.

Kaplan, S. & Kaplan, R. (1982) Cognition and Environment: Functioning in an Uncertain World (New York: Praeger).

Knoepp, J.D., Coleman, D.C., Crossley, D.A., Clark, J.S. (2000). Biological indices of soil quality: an ecosystem case study of their use. Forest Ecol. Manage. 138 (1-3), 357-368.

Koblet, O. and Purves, R.S., 2020. From online texts to Landscape Character Assessment: Collecting and analysing first-person landscape perception computationally. Landscape and Urban Planning, 197, p.103757.

Landers, D.H. and Nahlik, A.M., 2013. Final ecosystem goods and services classification system (FEGS-CS). Anonymous EPA United States Environmental Protection Agency. Report Number EPA/600/R-13/ORD-004914.

Luo, Y., Zhou, X. (2006). Soil respiration and the environment. Academic Press, London, Elsevier.

Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 2005; 33: 5691-702.

Pykälä, J. (2003). Effects of restoration with cattle grazing on plant species composition and richness of semi-natural grasslands. Biodivers. Conserv. 12 (11), 2211-2226.

Robinson, D.A., Thomas, A., Reinsch, S., Lebron, I., Feeney, C.J., Maskell, L.C., Wood, C.M., Seaton, F.M., Emmett, B.A. and Cosby, B.J., 2022. Analytical modelling of soil porosity and bulk density across the soil organic matter and land-use continuum. Scientific reports, 12(1), p.7085.

Schirpke, U., Hölzler, S., Leitinger, G., Bacher, M., Tappeiner, U. and Tasser, E., 2013. Can we model the scenic beauty of an alpine landscape?. Sustainability, 5(3), pp.1080-1094.

Spurgeon, D.J., Ricketts, H., Svendsen, C., Morgan, A.J. and Kille, P. (2005). Hierarchical responses of soil invertebrates (earthworms) to toxic metal stress. Environmental science & technology, 39(14), pp.5327-5334.

Tveit, M., Ode, Å. and Fry, G., 2006. Key concepts in a framework for analysing visual landscape character. Landscape research, 31(3), pp.229-255.

Wood SN. (2017). Generalized additive models: an introduction with R: CRC press, 2017.

Appendix 1

Table A1. Linking planned (or proposed) soil physico-chemical and other measures with NCPs in Soilguard (WP2, WP3, WP4 and WP5).

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
(Underpinning data)	texture	% sand, clay, silt	General soil characterization, not directly related to NCPs, but important to predict & interpret results	Upscaling (WP5): Derive from gridded datasets (via API, or pre-calculated): -May need stratification, for e.g. pH, bulk density, etc.	General methodological assumptions on valuation: -Assume steady state -Annual value vs Asset value -Calculate forestry over decadal scale, but reduce to annual value
	pH	Unitless			
	Electric conductivity	microS/cm			
	Bulk density	kg soil/m3	(measured in situ by local teams) Related to soil compaction. Within soil type, can be used as an estimate of soil compaction.	Check literature to classify these values into "compacted"	
Food and Feed	Yield	t/ha	Key variable. Challenge to represent in the app as general prediction, but possible for Soilguard data.	Options are: i) Data from study farms (from WP2) ii) Measure in experiments (from WP3) iii) Literature search on Organic vs Conventional Additional notes: Arable: Functions likely to be crop-specific. Grassland: Stocking density (cattle) from WP2 data, could reverse calculate grass productivity if needed. Forestry: Overall productivity available for different forest management types. Standardise for rotation period.	Market price; Rating Arable: Market prices needed separately for Conventional & Organic, by crop type (average over multiple years), also by country. Livestock: Additional steps needed to calculate value. Forestry: Market price for Timber

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
				Upscaling (WP5): Substantial challenge in the App.	
	Potential crop production			Agreed out of scope	
Climate regulation	Soil organic C	g. organic C/kg soil	Convert to carbon stock. Data available for WP2 sites. Extract from SoilGrids data using API	Stratify by land use type. Arable: Derive averages from WP2 data Grassland: As above Forestry: As above for soil C, Separately derive data on above- and below-ground C stocks for trees How to upscale (WP5): Stratify by land use type: Arable: Look up nearest C stock (or calculate from soil maps), assume this is for Conventional, apply factor to convert for organic. Grassland & Forestry. In principle, a similar approach to that for arable could apply.	Value change in C with shift of management. Carbon market price; Rating Social Cost of Carbon
Regulation of freshwater quantity (Flood regulation)	Water infiltration	Lab assay: Amount of time in infiltrating 50% of the 10 ml added	Challenging to relate to surface water flooding risk, but a high priority if possible.	Lab assay results are challenging to upscale. Alternative function from the literature (WP5).	Rating
	Water holding capacity	% (g. water retained/g. dry soil)	Helps with flood and drought regulation		

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
Regulation of freshwater quality	NO3/PO3 in leachates	Lab assay: mg N/ L of leachate, mg P/ L of leachate	Laboratory measure: column with 50 g of soil, water with 100 ml + a commercial fertiliser (known NO3/PO3).	May be possible to use this as a proxy measure for leaching risk. Upscaling (WP5): Substantial challenges	Rating
	Cation exchange capacity	CEC	Related to soil´s capacity to retain nutrients	CEC will not be measured.	
Soil formation and protection (Soil erosion prevention)	Soil aggregates stability	Semi-quantitative (scores 1 to 16)	Related to resistance to further erosion.	Combine with measure of overland flow risk to produce proxy metric of soil erosion risk Score can be related to Low, Med, High risk. Could benchmark values from WP2/WP3 data against mapped soil erosion risk. Upscaling (WP5): Challenging, but may be possible to relate to mapped soil erosion risk & apply factor for Organic vs Conventional. Incorporate slope in assessment. Assume only applicable to Arable	Costs of soil erosion; Rating Cost estimates are available, e.g. cost of dredging, cost of lost fertility, etc.
Soil formation and protection (Nutrient cycling)	Available P	mg P/kg soil	Plant-available nutrients	May be feasible to include in NCPs, as well as visualise in data-side of App, & could form basis of advice to farmers on options to improve at local level.	Replacement cost (of N fertiliser) (Unlikely to value in this context)
	Available N	mg NO3-/kg soil + mg NH4-/kg soil			
Nutrient cycling	Litter decomposition (tea bag index)	% weight loss/day		Multiple measures. Not feasible to include in NCPs, but could visualise in data-side of App (WP5)	Unlikely to value in this context

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
	Soil enzymatic activities (beta-glucosidase)	Micro-mols of p-nitrophenyl/ g. dry soil · hour	Potential capacity of the soil to degrade C sources		
	Soil enzymatic activities (acid phosphatase)	Micro-mols of p-nitrophenyl/ g. dry soil · hour	Potential capacity of the soil to obtain P		
	Potential N mineralization	units	Potential capacity of the soil to transform ammonia to nitrate		
	Degradation C sources	Microresp (units)	Potential capacity of the soil to degrade multiple C sources		
	N cycle genes	AmoA/AmoB, INRAE is doing those	Quantifies different pathways of the N cycle		
	Amount of mycorrhizal fungi	% of taxa/taxa in each sampled, NLFAs (UvA)	Mycorrhizal fungi aids crop growth, mainly under nutrient-limiting conditions		
Regulation of detrimental organisms (Biological control)	Leaf damage	% of leaf surface damaged by pathogenic fungi and%	Not available for some sites.	Leaf damage difficult to take forward in meaningful way for NCPs.	Replacement cost (of pesticides); Rating (May be possible. Further assumptions then needed on how to relate reduction in yield to typical

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
		of leaf surface damaged by herbivorous insects			cost of pesticides in conventional systems & any associated costs for biological control in organic systems) **Substantial work needed
	Amount of root feeding nematodes and other invertebrate pests	% of taxa/taxa in each sample		Proportion of root-feeding nematodes, and proportion of microbial pest-types could both be expressed in a form which gives proxy for regulating service. Requires assumptions on level of pressure related to threshold level of abundance of pest organisms. May require literature review, or benchmark against yield data, or compare against yield gap. **Substantial work needed [?] Applicable to Arable, Grassland & Forestry. Upscaling (WP5): Potentially possible via link to function/composition but this aspect currently uncertain.	
	Amount of microbial pests	% of taxa/taxa in each sample that are known pests for crops			
	Antibiotic resistance genes			Agreed out of scope	
Habitat creation	Hedgerows and treelines	% area covered by these linear green structures	Related to habitat provision. (200m, 500m and 1km radius around each site). From Copernicus (2014-2017), for EU Regions.	Can combine measure of habitat extent with a measure of connectivity (WP2 & WP5).	Rating. Unsure if other measures available. (Avoids double counting with aesthetics below if include connectivity)
	Naturalness	% area covered by	Related to habitat provision. (200m, 500m and 1km radius	Agreed out of scope	

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
		natural habitat	around each site). From Copernicus (2014-2017)		
	Geodiversity	Unknown units.	Related to habitat heterogeneity for soil organisms. Diversity of soil types and geomorphological features (200m, 500m and 1km radius around each site).	Agreed out of scope	
Physical and psychological experience (Aesthetic landscapes)	Hedgerows and treelines	% area covered by these linear green structures	Related to habitat provision. (200m, 500m and 1km radius around each site). From Copernicus (2014-2017).	Could be combined with some other metrics, e.g. field size, water. Upscaling (WP2 & WP5): Would need separate pre-calculated measures for all Europe – can't be done Live in App.	Rating Unsure if other measures available
	Field size		Data from WP2 surveys, and remote sensing		
Physical and psychological experience (Aesthetic landscapes)	Diversity of birds	Spp. abundance	Per-region species lists, obtained from eBird (citizen science), GBIF etc.	Challenging to create an informative metric for NCPs.	Rating Unsure if other measures available
	Diversity of plants	Spp. abundance		Upscaling (WP5): Challenging, but could use Conventional vs Organic factors, and assume Conventional as default.	
	Number of ecotourism companies nearby			Agreed out of scope	
	Number of biodiversity studies in the region			Agreed out of scope	

Related NCP	Measurement	Units	Comments	Notes on additional work	Valuation (WP4)
	Number of visitors in the region			Agreed out of scope	