Ref. Ares(2024)3555461 - 16/05/2024

# Roadmap for research, capacity and financing options for scaling up BESTMAP approach

## Deliverable D5.5

31 January 2024

Guy Ziv[1], Catriona Willoughby[1], Denise Hick[1], Arjan Gosal[1], Rosemary Wool[1], Jenny Li[1], Jiaqi Ge[1], Paul M. Evans[2], Michael Beckmann[3], Anne Paulus[3], Meike Will[3], Birgit Müller[3], Tomáš Václavík[4], Tomáš Čejka[4], Ruth Delzeit[5], Ornella Dellaccio[6], Jon Stenning[6], Anna Cord[7], Stephanie Roilo[7], Sanja Brdar[8], Joan Maso[9], Yoni Gavish[10], Danny Hooftman[11], James M. Bullock[2]

[1]*University of Leeds*
[2]*UK Centre for Ecology and Hydrology*
[3]*UFZ Helmholtz Centre for Environmental Research*
[4]*Palacký University Olomouc*
[5]*University of Basel*
[6]*Cambridge Econometrics*
[7]*Technical University of Dresden*
[8]*Biosense Institute*
[9]*Centre for Ecological Research and Forestry Applications*
[10]*Data-Science-Environment: Statistical and Ecological Solutions*
[11]*Lactuca*

**BESTMAP**
**Behavioural, Ecological and Socio-economic Tools for Modelling Agricultural Policy**

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817501.

**Prepared under contract from the European Commission**

Grant agreement No. 817501
EU Horizon 2020 Research and Innovation action

| | |
|---|---|
| Project acronym: | **BESTMAP** |
| Project full title: | **Behavioural, Ecological and Socio-economic Tools for Modelling Agricultural Policy** |
| Start of the project: | September 2019 |
| Duration: | 54 months |
| Project coordinator: | Prof. Guy Ziv<br>School of Geography,<br>University of Leeds, UK<br>[http://bestmap.eu/](http://bestmap.eu/) |
| Deliverable title: for | Roadmap for research, capacity and financing options<br>scaling-up BESTMAP approach |
| Deliverable n°: | D5.5 |
| Nature of the deliverable: | Report |
| Dissemination level: | Public |
| WP responsible: | WP5 |
| Lead beneficiary: | University of Leeds |
| Citation: | Ziv, G., Willoughby, C., Hick, D. Gosal, A., Wool, R., Li, J., Ge, J., Evans, P.M., Beckmann, M., Paulus, A., Will, M., Müller, B., Václavík, T., Čejka, T., Delzeit, R., Dellaccio. O., Stenning. J., Cord, A., Roilo, S., Brdar, S., Maso, J., Gavish, Y., Hooftman, D. and Bullock, J.M., (2024). *Roadmap for research, capacity and financing options for scaling up BESTMAP approach*. Deliverable D5.5 EU Horizon 2020 BESTMAP Project, Grant agreement No. 817501. |
| Due date of deliverable: | Month n°53 |
| Actual submission date: | Month n°53 |

Deliverable status:

| Version | Status | Date | Author(s) |
|---|---|---|---|
| 1.0 | Draft | 31 January 2024 | Guy Ziv, Catriona Willoughby, Denise Hick, Arjan Gosal, Rosemary Wool, Jenny Li, Jiaqi Ge, *University of Leeds*<br>James M. Bullock, Paul M. Evans, *UK Centre for Ecology and Hydrology*<br>Michael Beckmann, Anne Paulus, Meike Will, Birgit Müller, *UFZ Helmholtz Centre for* |

*Environmental Research*
Tomáš Václavík, Tomáš Čejka, *Palacký University Olomouc*
Ruth Delzeit, *University of Basel*
Ornella Dellaccio, Jon Stenning, *Cambridge Econometrics*
Anna Cord, Stephanie Roilo, *Technical University of Dresden*
Sanja Brdar, *Biosense Institute*
Joan Maso, *CREAF*
Yoni Gavish, *Data-Science-Environment: Statistical and Ecological Solutions*
Danny Hooftman, *Lactuca*

# Table of contents

## Summary

This Deliverable provides a roadmap to expansion of BESTMAP towards a operational pan-European modelling platform, as well as explore via pilot analyses several areas for improvement and future research.

Considering new case studies, we analyse the locations where models parameterized in those regions can transfer to cover the most area. We conclude that in future case studies, they should be located in northern Spain, north-west Italy, central Italy, Montenegro/Albania, and Bulgaria.

Testing if one can model water quality at the European scale, our modelling shows the NDR model (used in BESTMAP CS work) has generally good performance at EU scale, despite it being a rather simple process-based model. There is an overestimation of Nitrogen at low N, and underestimation of Phosphate at high P, which need to be considered in future work.

Regarding soil organic carbon (SOC), we performed a more in-depth literature review and summarised impacts of different management practices on SOC stocks. Our work has identified several gaps in the current body of literature that should be addressed by future research. Our analysis showed that much of our knowledge on soil carbon stock change is limited to the top 50 cm of the soil, there is a need for a development in our understanding of combinations of management practices and how they may impact carbon stocks in the field, and a clear need for sharing of information long-term experiments to assess long-term storage and “permanence” of SOC.

We also explored if SOC can be determined from remote sensing. Specifically, this pilot analysis looked at extracting the spectral signature of bare soil from Sentinel-2, using LUCAS sampled bare agricultural fields. We explored the spectral changes with soil moisture of one UK soil, but extending that to other EU soils is needed to build synthetically dry spectra over the whole of Europe.

Via an expert workshop entitled “ABM/CGE modelling” on May 12-13 2022 in Basel, we explored linking CGE/PE models to ABM. We discussed potential issues, specifically focused around data availability and conceptual alignment of those modelling approaches. The agenda for the workshop can be found in the Appendix.

Focusing more on data, we discuss the potential for improving access to data, in particular from Farm Accountancy Data Network (FADN). We discuss issues around usefulness, access, compatibility and sampling. We also performed a pilot testing if synthetic data generated from FADN can be useful for future modelling efforts following the BESTMAP approach.

We close with a discussion around financial options to extend the work BESTMAP has conducted in the five case studies.

## Preface

This Deliverable 5.5 of the BESTMAP project complies with the activity of Task 5.4 of the Grant Agreement, namely to “...create a white paper roadmap reviewing the research, regulatory, capacity and financing requirements to scale up the approach of BESTMAP to be used as an operational policy impact assessment platform by the EC".

The Conceptual Framework of the BESTMAP-Policy Impact Assessment Model is visualised in the following Figure 1, taken from D2.5 "BESTMAP Conceptual Framework Design & Architecture (update)". In brief - spatially explicit biophysical and socio-economic models were developed in (5) case studies - for biodiversity (farmland birds), water quality, soil carbon, food production and net farm added value (see D3.3). These were based on characterising individual farms into archetypes (see D3.5) and used data from IACS/LPIS. In addition, agent based models (ABMs) were parameterized to explain the adoption of agri-environmental schemes (AES) by individual farms (see D4.1). These outputs of these models were used to parameterize farm-level regression models (see D5.2) and a novel approach to assess how "far" (in a multi-dimensional space) one can transfer those models across FADN regions was developed (see D5.1). All outputs were translated into policy-relevant indicators (see D4.3) and put into an online dashboard (see D2.3) and policy notes (see D6.7).

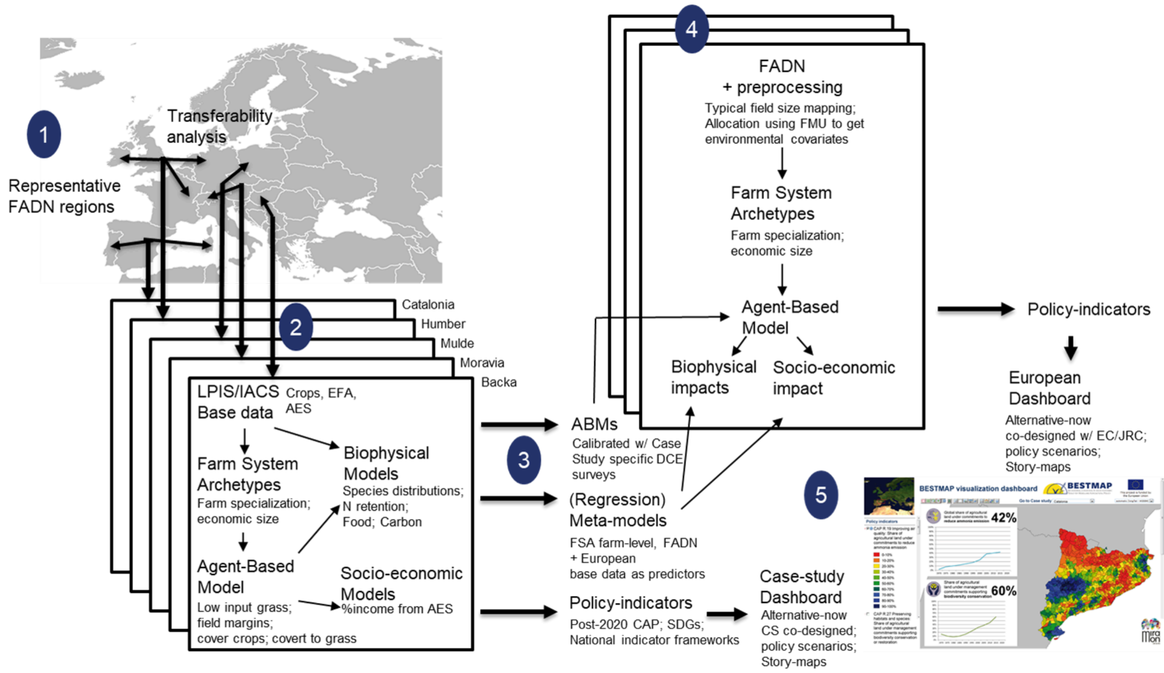


***Figure 1:*** *Overview of the BESTMAP-Policy Impact Assessment Model (BESTMAP-PIAM) framework*

The aim of this Deliverable is to explore some aspects of extending and improving on the work performed by the BESTMAP project. In particular:

1. Where should new Case Studies be located to allow transferability of models to areas not currently covered by the case studies in BESTMAP? (Section 1)
2. Is it possible to find the data and run those biophysical models, exemplifying with the Nutrient Delivery Ratio for water quality, across the whole of Europe (Section 2)

3. Can we improve on some key models, specifically the soil carbon model, and what will be needed for that across Europe? (Section 3)
4. Given the importance of soil carbon as a determinant of soil quality, which was used by the ABMs for allocation of AES, can we do better prediction of SOC using Remote Sensing data, extending on Task 5.3 activities? (Section 4)
5. What way can future projects link ABMs with more traditional macroeconomic models? (Section 5)
6. What are some data issues, barriers and potentials for future projects? (Section 6)
7. With the importance of FADN, but also the sensitivity and difficulty working with it, can synthetic data generated from FADN be used instead of the original data? (Section 7)
8. What may be some financial options to extend the work of BESTMAP project (Section 8)?

# 1. Scaling up the BESTMAP approach

The upscaling phase in BESTMAP involved generating meta-models for ecosystem services in NUTS3 regions within each case study area. These meta-models utilised the case study ecosystem service model result as the response variable, while initial sets of potential explanatory factors were determined from environmental and economic predictors based on expert opinion. Variable selection techniques were employed to streamline the explanatory variables, resulting in a distinct set of predictors for each ecosystem service per NUTS3 region. Each meta-model was used to predict the results of each NUTS3 region within any of the case study areas, with the coefficient of determination ($R^2$) indicating how well each model could predict. The resulting $R^2$ values were plotted against the Minkowski distance (i.e. a metric of similarity) of the initial group of possible explanatory factors per ESS. This enabled the upscaling team to determine the 'transferability' of ESS results to other non-case study NUTS3 regions across the whole of Europe. If plotted results met certain 'transferability criteria', then we had high enough confidence in those results to be able to transfer them to new regions based on their Minkowski distance. For further details on the above methodology or the results, see Deliverables 5.1 and 5.2.

Some NUTS3 regions within Europe (which, for this work, included those in Turkey, as they had been assigned NUTS3 regions) could not be well-predicted using the above methodology. This was largely because the Minkowski distances between those regions and the current case studies were too large, meaning that if the BESTMAP work was to be repeated in the future other case studies in different locations would be required. To keep resource expenditure as low as possible, any new case studies would have to represent (i.e. be close in Minkowski distance to) as many regions as possible that are not satisfactorily covered in the BESTMAP project. Therefore, in this section, we sought to identify potential locations and the necessary number of representative future case studies for new case studies if the actions of BESTMAP were to be repeated.

The methodology was split into two parts. The first involved establishing the representativeness of all NUTS3 regions in relation to each other. This assessment was grounded in the predictors employed during the upscaling analysis of ESS provision. The second involved identifying regions where confidence levels were inadequate for the transfer of ecosystem service outcomes, based on the existing BESTMAP case studies. Note: NUTS3 regions >700 km away from 'mainland' Europe were excluded from this analysis.

### Representativeness of all NUTS3 regions in relation to each other

To identify NUT3 regions that represented other NUT3 regions well, we used hierarchical clustering. This technique allowed similar data points (in this case, NUTS3 regions) or objects to be sorted into clusters or groups based on their similarities across multiple variables. Each data point starts as a separate cluster and is progressively merged into larger clusters, finally being represented as a tree-like structure (i.e. dendrogram). The dendrograms indicate which regions were most similar to each other, while also giving an indication of the relationship between all of the regions.

The choice of which distance matrix is used is important for hierarchical clustering, and has a major impact on the results of the clustering. Here, we used the same logic that was used for the Upscaling part of BESTMAP, namely that the environmental predictors of an ESS that were considered important for the ESS provision were included (see D5.2). Therefore, the distance matrices were different for different ESS. These individual matrices were calculated as part of D5.2 and were taken directly from that output.

Another important consideration when using hierarchical clustering is from where the distance is calculated on the dendrograms (i.e. the height at which a dendrogram is 'cut'), which leads to different numbers of resulting cluster groups. This has the advantage over other forms of clustering (e.g. k-means clustering) due to the number of clusters not having to be defined beforehand. To ensure consistency, a single distance value (height on a dendrogram) was used for all ESS clustering. This height had to be relevant to all of the NUTS3 regions across Europe, and to all of the ESS; therefore, a summary measure of all four distance matrices was used. The distribution of each 'lower triangle' of each ESS distance matrix (i.e. entries below the main diagonal) was obtained, from which the median was calculated (Figure 1.1). The mean of these four medians was then computed, giving a final dendrogram cut height of 3.76.

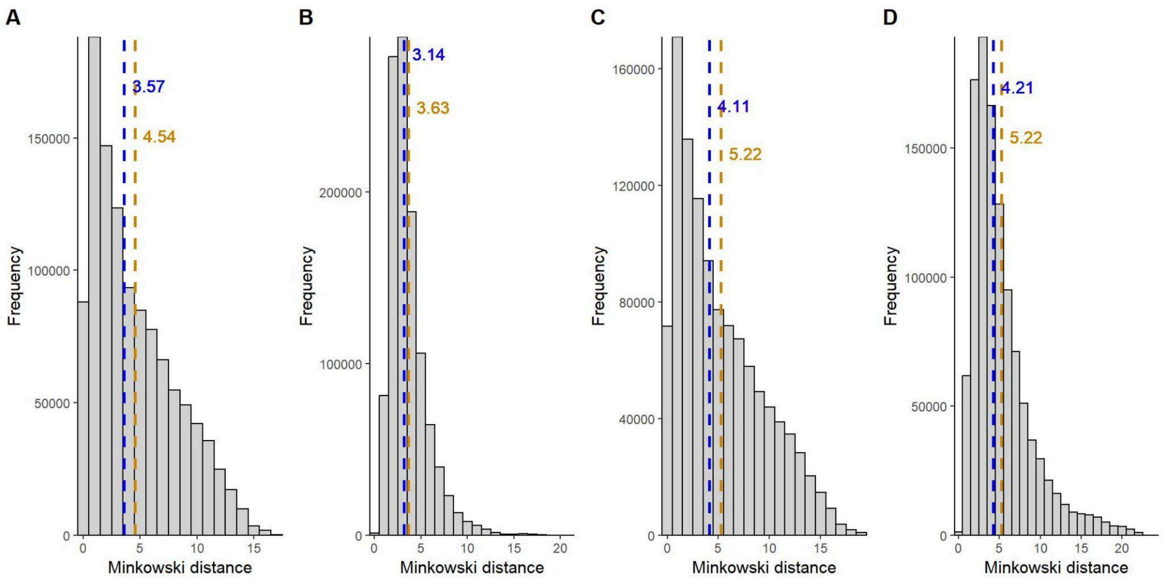


***Figure 1.1:*** *Histograms of Minkowski distance metrics between all NUTS3 regions for different ESS: A = biodiversity; B = carbon; C = food; and D = nutrient. Vertical dashed lines represent medians (blue) and means (orange).*

We used the calculated cut height to determine the final cluster groups for each ESS. This resulted in the four ESS being split into 10, 80, 42, and eight clusters for food, nutrients, carbon, and biodiversity, respectively. Food (Figs 1.2 and 1.3) and biodiversity (Figs 1.8 and 1.9) both had fewer final clusters due to the fewer variables that were considered important driving factors for those ESS (and therefore went into the distance metrics calculations) including different aspects of climatic data. Climatic data, especially across Europe, generally has less variability along the same latitude due to the moderating influence of large water bodies, such as the Atlantic Ocean, which helps regulate temperatures and create more stable weather patterns. Conversely, climate factors such as precipitation and temperature can alter markedly with changing longitudes. Combined, these create some very distinct regions of Europe, with the regions not being particularly similar in multi-dimensional climate space (Fig 1.3).

The nutrient (Figs 1.4 and 1.5) and carbon (Figs 1.6 and 1.7) clustering resulted in many more final classes. The driving factors for these ESS were likely to be less affected by climatic phenomena (although temperature and precipitation predictors were still included for both) and more by geo-physical data; predictors included soil pH, soil bulk density, soil clay content, topsoil carbon, soil moisture, and soil type for carbon, and soil clay content, small woody features, soil moisture, elevation, and soil type for nutrient export. None of these extra data groups form such distinct spatial patterns across the whole of Europe, especially at a relatively coarse (i.e. NUTS3-level) scale. For this reason, there is generally less similarity between clusters that include such data.

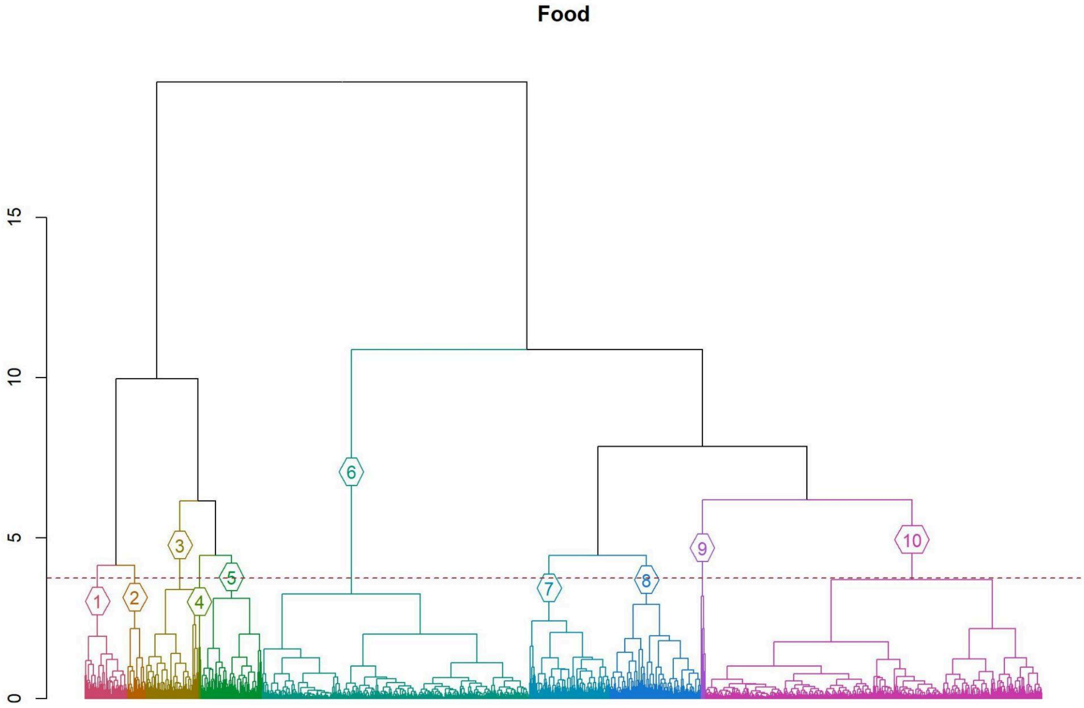


***Figure 1.2:*** *Final clusters based on Minkowski distance of environmental predictors for food production. Each cluster is represented by a different colour and its number is indicated inside the hexagon. The brown dashed line is the cut height.*

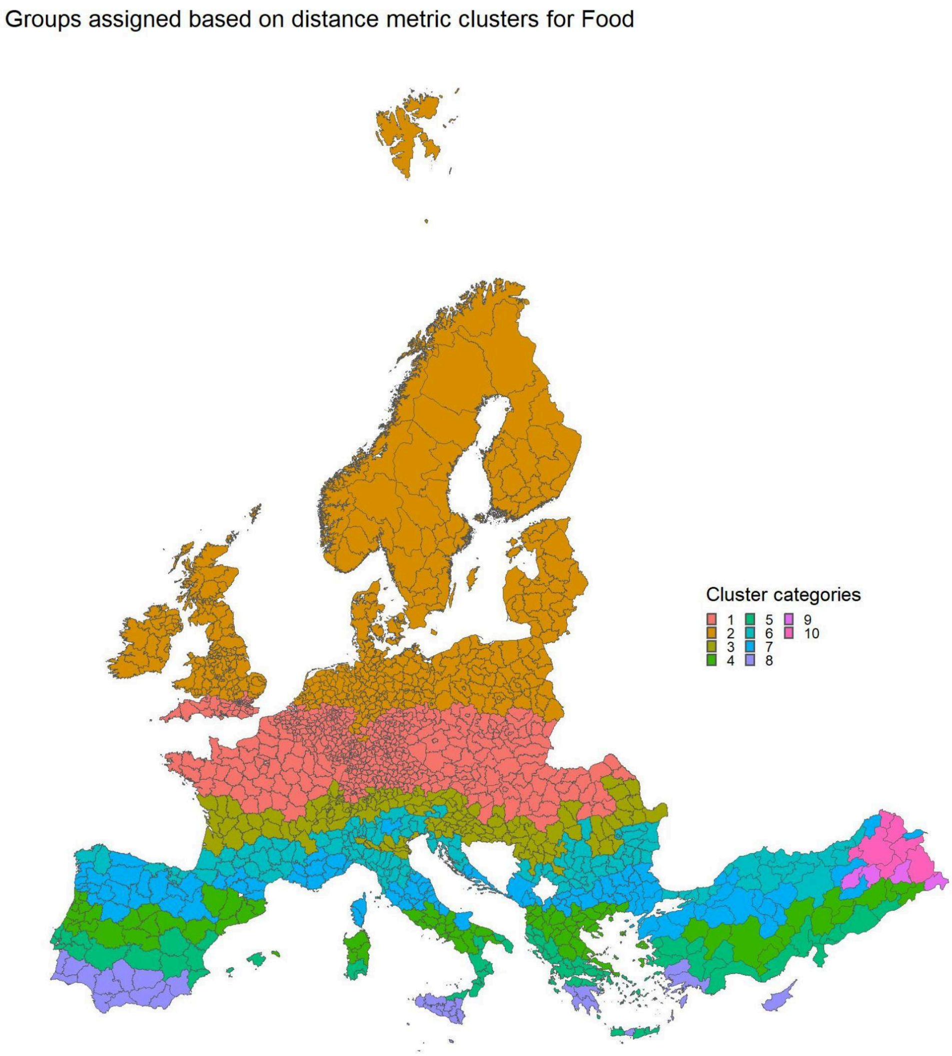


***Figure 1.3:*** *Final clusters based on Minkowski distance of environmental predictors for food production, represented spatially.*

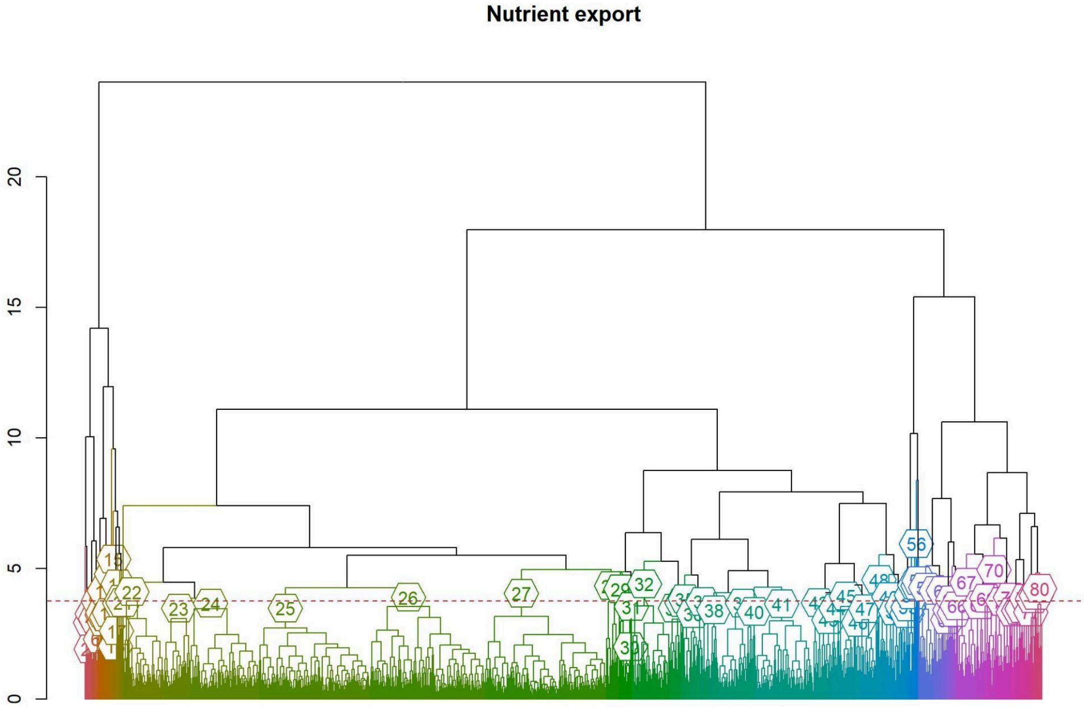


***Figure 1.4:*** *Final clusters based on Minkowski distance of environmental predictors for nutrient export. Each cluster is represented by a different colour and its number is indicated inside the hexagon. The brown dashed line is the cut height.*

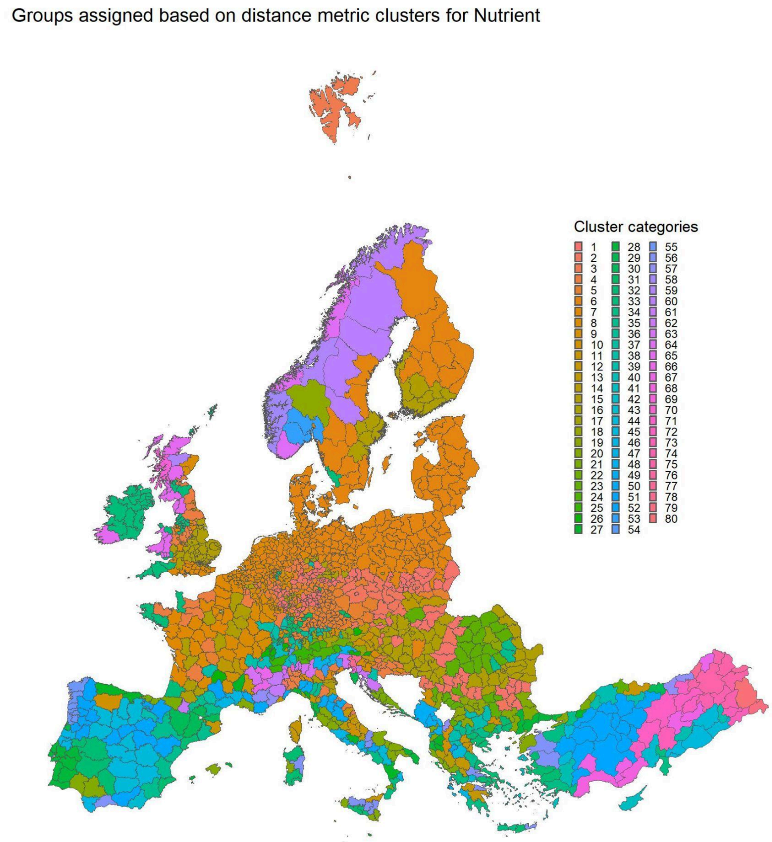

***Figure 1.5:*** *Final clusters based on Minkowski distance of environmental predictors for nutrient export, represented spatially.*

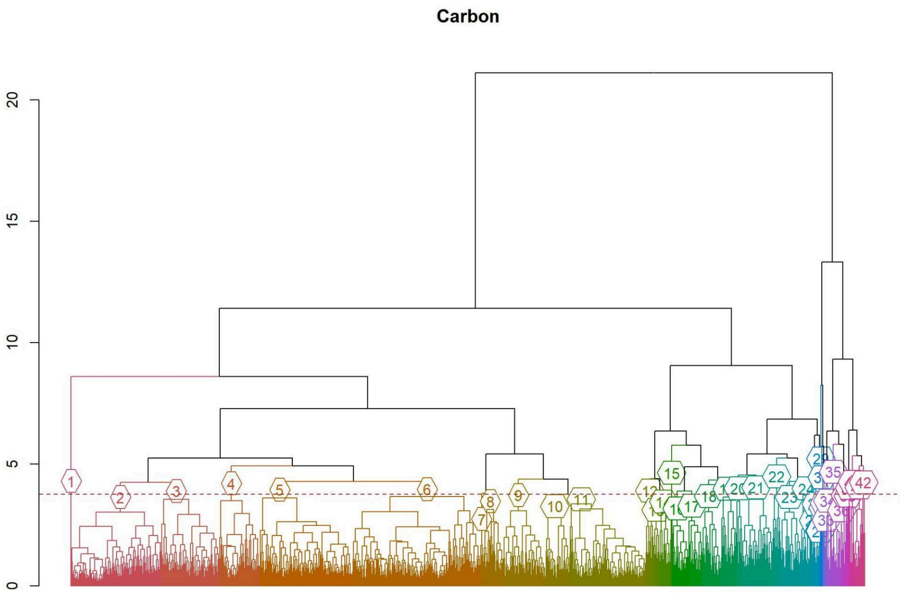


***Figure 1.6:*** *Final clusters based on Minkowski distance of environmental predictors for carbon. Each cluster is represented by a different colour and its number is indicated inside the hexagon. The brown dashed line is the cut height.*

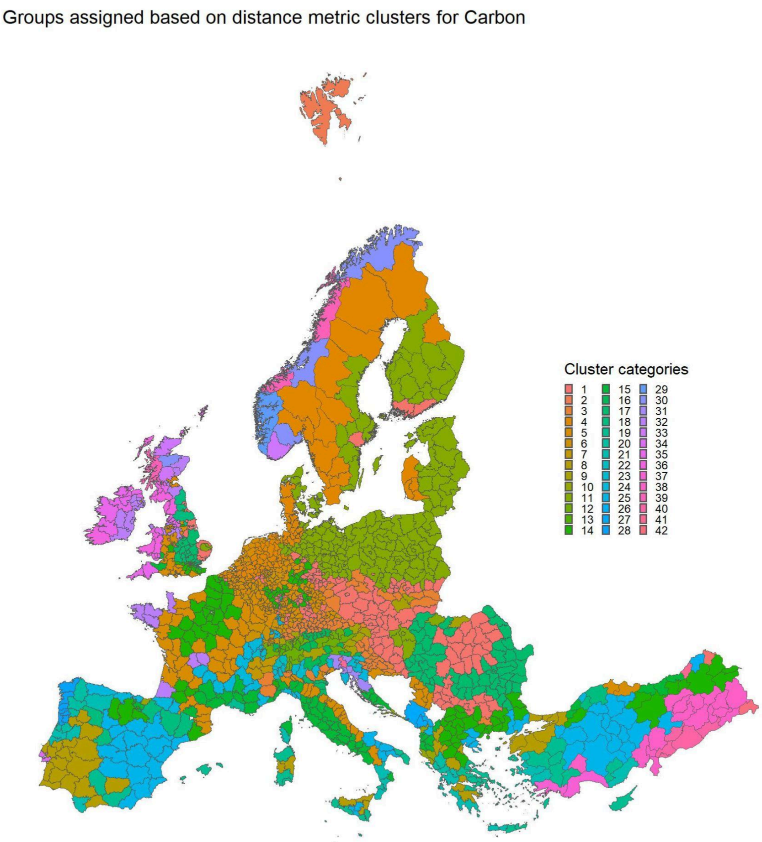


***Figure 1.7:*** *Final clusters based on Minkowski distance of environmental predictors for*

_____________________________________________________________________________

*carbon, represented spatially.*

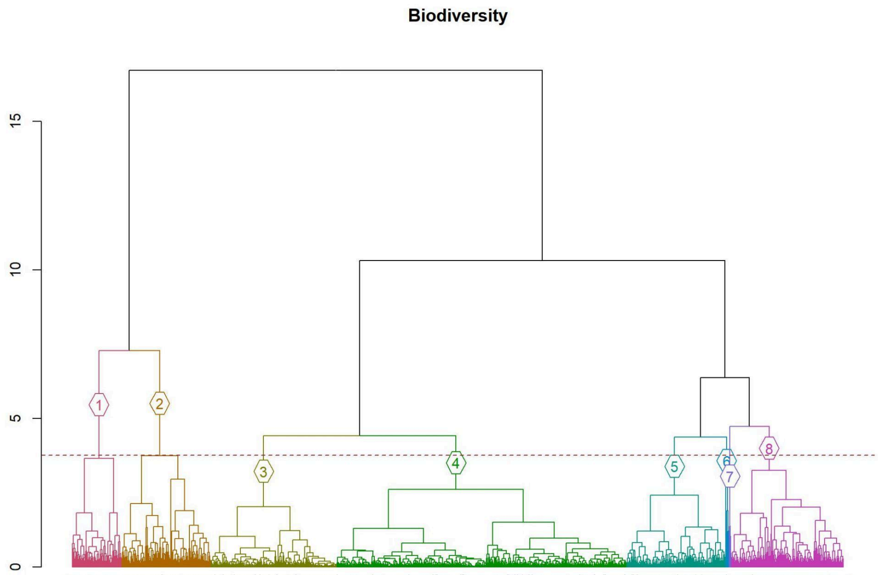


***Figure 1.8:*** *Final clusters based on Minkowski distance of environmental predictors for biodiversity. Each cluster is represented by a different colour and its number is indicated inside the hexagon. The brown dashed line is the cut height.*


***Figure 1.9:*** *Final clusters based on Minkowski distance of environmental predictors for*

*biodiversity, represented spatially.*

Since the BESTMAP work focused on all four ESS, it is important that any future case studies represent areas that are sufficiently representative of all four ESS in terms of their possible driving predictors. Therefore, we combined the four cluster map outputs from the four ESS (Figures 1.3, 1.5, 1.7, and 1.9). To do this, the four ESS cluster maps were overlaid, which created a unique combination of the original four ESS clusters. For example, cluster 1 in Fig. 1.10 consisted of the cluster 1 of biodiversity, cluster 1 of food, cluster 1 of nutrient, and cluster 12 of carbon (Figure 1.10). These combinations resulted in many, 440, combinations, after the current case study regions were excluded (Figure 1.10).

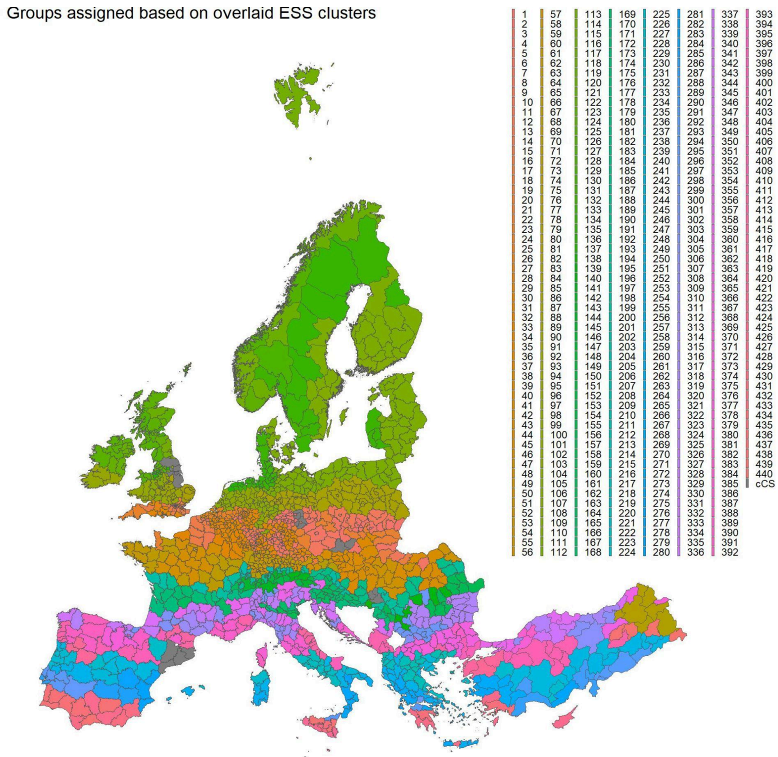


***Figure 1.10:*** *Unique cluster combinations resulting from overlaying all four ESS maps and their NUTS3 cluster classifications. The dark grey areas represent BESTMAP's current case studies (labelled 'cCS').*

## Identifying regions for future case studies

The location of the most useful future case studies required the identification of areas that are not currently sufficiently covered, as determined by the transferability criteria. To identify such regions, we excluded any NUTS3 region that met the transferability criteria more than five times for any one of the ESS. For example, if the transferability criteria were met in a NUTS3 region 7, 6, 2, and 1 times (one for each ESS), that NUTS3 region would be excluded. However, if the criteria were met 5, 4, 2, or 1 times it would be included.

The resulting map was reduced to 166 cluster combinations (Figure 1.11). It highlighted different regions where transferability confidence is low based on the current BESTMAP case studies. These regions indicate where future case studies could be based (in addition to the current five BESTMAP CS regions) in order to be able to cover all of Europe in terms of transferring ESS models run at the case study level.

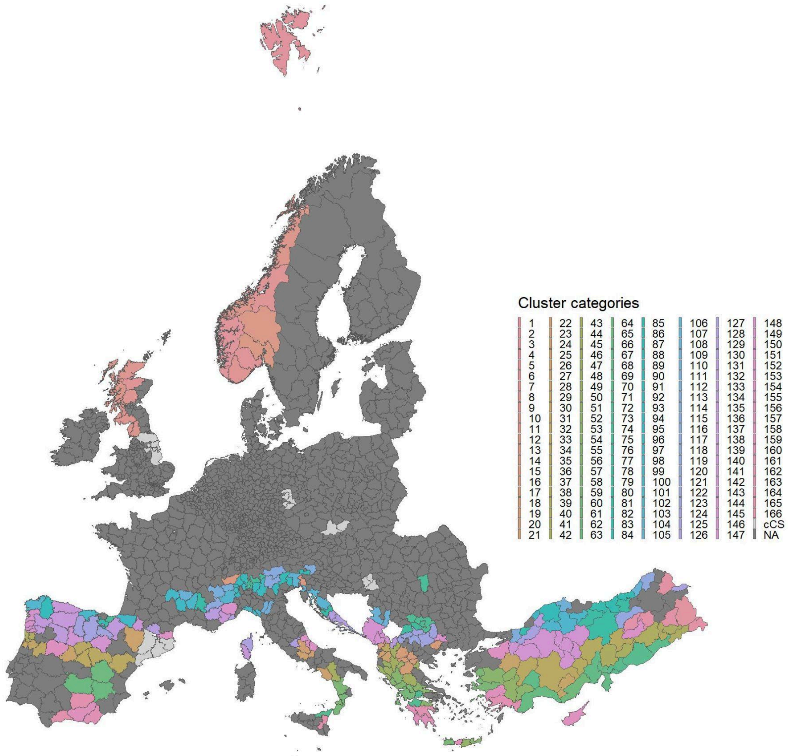


***Figure 1.11:*** *Cluster combinations that remain after any NUTS3 regions that met the transferability threshold criteria at more than 5 times per ESS were excluded. The dark grey areas represent areas that were removed (labelled 'NA') because they were sufficiently covered by existing BESTMAP case studies; and the light grey areas represent BESTMAP's current case studies (labelled 'cCS').*

As there was funding for five case study regions in the BESTMAP project, it seems reasonable to assume that a similar amount may be funded for another project. Therefore, to determine a more defined shortlist of five or fewer suitable places, a more refined criterion was used: we excluded any NUTS3 region that met the transferability criteria more than **three** times for any one of the ESS, which gave fewer but more-distinct regions. Additionally, we excluded any based in Turkey, due to that country not being a current EU Member State . By conducting this further analysis, we identified regions that align with the current project's goals and have the potential for successful implementation (Figure 1.12). We conclude that to get the optimal value from future case studies, they should be located in northern Spain, north-west Italy, central Italy, Montenegro/Albania, and Bulgaria.

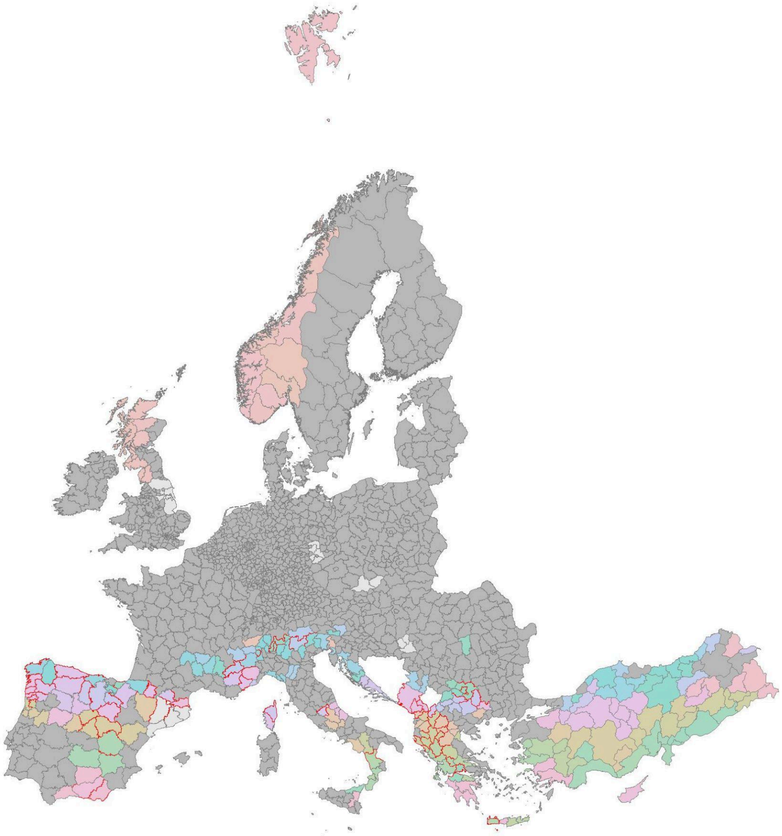

***Figure 1.12:*** *Cluster combinations that remain after any NUTS3 regions that met the transferability threshold criteria at more than* ***five*** *times per ESS were excluded (non-grey regions; see Fig. UP11), and those that remain after any NUTS3 regions that met the transferability threshold criteria at more than* ***three*** *times per ESS and NUTS3 regions in Turkey were excluded (regions outlined in red). The latter of these areas represent locations where additional case studies should be placed in addition to the current BESTMAP case studies to ensure transferability coverage of ESS models for the entirety of Europe. The dark grey areas represent areas that were removed (labelled 'NA') because they were sufficiently covered by existing BESTMAP case studies; and the light grey areas represent BESTMAP's current case studies (labelled 'cCS').*

As a final note, we want to highlight that the methodology described in this section is only the first step in identifying new case study regions; it is also crucial to engage with local stakeholders, including government officials, and communities, to gain insights into the unique challenges and opportunities each region presents. Challenges may include access to appropriate data at a regional level (e.g. LPIS (Land Parcel Identification System) data), which was a major issue within BESTMAP. Ultimately, the goal is to select regions that not only meet project criteria but also have the potential for long-term impact and sustainability.

## 2. Development and validation of a Europe-wide nutrient run-off model

In BESTMAP, we developed bespoke ecosystem service (ESS) models for each case study (Deliverable 3.3). To scale these up and enhance policy relevance, we developed approaches to assess transferability of these local ESS models to all other parts of Europe (Deliverables 5.1, 5.2). Here we assess an alternative approach; applying a single model across Europe, using Europe-wide data sources to parameterise the model. Specifically, we carry out a Europe-wide application of the InVEST Nutrient Delivery Ratio model (NDR) for estimating run-off of nitrogen (N) and phosphorus (P) from agricultural land. This NDR model was also applied in the case studies, and Deliverable 3.3 gives details of the model and the case study findings. In this section, we report our parameterisation of this model for N and P, which are important pollutants of waterways, and validate model predictions against measurements of N and P in European rivers at 2,252 locations. The workflow is summarised in Figure 2.1.

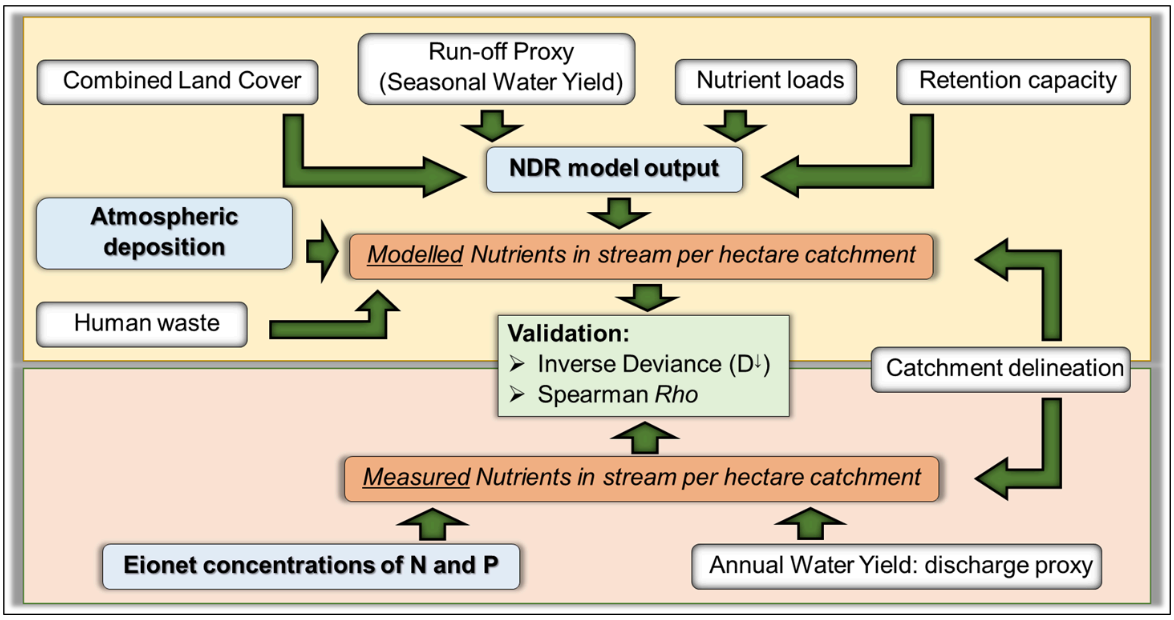


***Figure 2.1:*** *The workflow for running and validating the InVEST Nutrient Delivery Ratio model*

### Parameterising the NDR across Europe

The area modelled comprised the whole of the European Union along with the associated countries the United Kingdom, Switzerland, Norway and Iceland, as well as the countries comprising the former Yugoslavia. Country definitions used follow FAO (2015). While not shown in the figures, parts of Belarus and the Ukraine were modelled to encompass cross border catchments. The modelled scale was 25m within the GRS-1980-IUGG-1980 Lambert Azimuthal Equal Area projection (EPSG 9820). All inputs were projected to EPSG 9820 and resampled to the exact gridcell size and extent as the combined Land Cover map that we used.

The InVEST NDR model requires the following inputs, which were created using the datasets listed in Table 2.1: digital elevation map, land cover map, nutrient run-off proxy, the loads of N and P per vegetation class, the retention capacity per vegetation class, and atmospheric deposition and human waste. The approaches were as standard, but our calculation of the retention capacity per vegetation class deserves more detail.

For the retention capacity of the vegetation, studies as Redhead et al. (2018), Zawadzka et al. (2019) and Lavorel et al. (2022) were restricted in their sources and in the variation among vegetation types. Here, we explored a spatial approach, relating the amount of retention directly to the amount of Normalised Difference Vegetation Index (NDVI), which is a measure of the amount of vegetation. Using NDVI values from the Modis Terra satellite data (Table 2.1) we calculated retention for the 41 vegetation classes of the Land Cover map as:

$$\mathit{Retention\%} = \alpha(\mathit{NDVI}\text{-}\beta)^2$$

To ensure similar values to in Redhead et al. (2018) and Zawadzka et al. (2019), we needed the minimum median NDVI for woodlands to equate to 0.9, bare areas to 0.05, and croplands around 0.25. To achieve this, we set $\alpha$ to 11 and $\beta$ to 0.3, the latter being the minimum within-class median NDVI. Since 100% and 0% retention is only theoretically possible, vegetation classes values above 0.9 were set to 0.9 – a maximum retention value of 90%; values below 0.05 were set at 0.05, except for water (0).

| **Dataset Name** | **Year- Resolution** | **Detail & Usage** |
|---|---|---|
| **Spatial datasets used to generate the Land Cover map** | | |
| The EU crop map (d' Andrimont *et al*. 2021) | 2018 – 10m | Base map for assigning land cover |
| ESA WorldCover 2020 (ESA 2020) | 2020 – 10m | Map to fill the land covers not covered by the EU Crop map |
| EEA - Concentrations of nitrogen and phosphorus in European agricultural soils (EEA 2020) | 2019 – 1km | Refined non-improved grassland definition |
| Robinson *et al*. (2014): Livestock distribution | 2014 – 1km | LSU densities to subdivide grasslands |
| **Input datasets for the InVest Nutrient Retention Model and Water Yield Model** | | |
| Mineral Fertilizer Use per Crop (Ludemann *et al*. 2022) | 2018 – Country | Average application of N and P fertilizer per crop type |
| Manure Fertilizer use per country, EuropeAgriDB v1.0 database (Einarsson *et al.* 2021) | 2018 – Country | Average application and excreted of manure in grasslands and croplands |
| EEA – Nitrogen Use Efficiency (NUE) and Phosphorus Use Efficiency (PUE) EEA (2020) | 2019 – 1km | Concentrations of nitrogen and phosphorus in European agricultural soils |
| Global Hydrologic Soil Groups (HYSOGs250m; Ross *et al*. 2018) | 2020 – 250m | Hydrological soil groups |
| Slopes from EU-DEM v1.1 | 2016 – 25m | Five slope classes |
| Country ID (FAO 2015) | 2014 – Country | Countries following the FAO Global Administrative Unit Layers for 2014 |

| | | |
|---|---|---|
| EU-DEM v1.1 (Copernicus 2016) | 2016 – 25m | Combined SRTM and ASTER GDEM. |
| Terra NDVI for soil retention (MOD13Q1 v061; MODIS 2023a) | 2018 – 250m | 16-day NDVI calculated into retention capacity per Land Cover category |
| HydroSHEDS v1 – European Catchments (Lehner *et al*. 2009) | 2016 – ≈90m | Catchment definitions |
| Ecofloristic zones (Ruesch & Gibbs 2008) | 2006 – shapes | Ecofloristic zones used to subdivide the Land Cover Map into functional units |
| Terra Leaf Area Index (MOD15A2H v061; MODIS 2023b) | 2018 – 500m | 8-day LAI, recalculated and averaged to per month; used to calculated Kc-values |
| Terra Potential Evapotranspiration (MOD16A2 v061; MODIS 2023c) | 2018 – 500m | 8-day Evapotranspiration |
| WorldClim v2.1. Precipitation (Fick & Hijmans 2017) | Period – ≈1km | Historic average climatic precipitation per month (1970-2000) |
| FAO Wet days frequency (New *et al*. 2002) | Period – ≈10km | Historic average rain days per month (1961-1990) |
| Global Hydrologic Soil Groups (HYSOGs250m) (Ross *et al*. 2018) | 2020 – 250m | Soil Groups for Curve Number (CN)-based runoff modelling |
| Soil Depth from Hooftman *et al*. (2023) | 2013 – ≈1km | Soil Depth per vegetation class |
| ESDAC Available Water Content (Ballabio *et al*. 2016) | 2009 – 500m | Topsoil physical properties for Europe |
| USDA (2004) | Non spatial | CN values per soil group |

***Table 2.1:*** *Datasets used to run a Europe-wide NDR model*

## Model validation

Model output was validated against within-river nutrient concentrations measured under the European Environment Information and Observation Network (Eionet), a partnership network of the European Environment Agency (Wise v6, EEA 2022). We selected those measurement locations from the aggregated data which had more than a mean of 10 measurements of N and more than 15 for P after 1995. N was a combination of $NO_x$ and $NH_x$, with their atomic weights recalculated into sole N.

The data are in concentrations, mg/l, so we required an annual river discharge estimate to calculate the annual amount of nutrients passing through the given points in rivers to match the modelled units. We generated annual river discharge using the InVEST Annual Water Yield model (AWY), using inputs from Table 2.1. The AWY is mathematically dissimilar from the Seasonal Water Yield model used for the NDR calculations as it is not stream network based and some of the input data are different. Therefore, validation data and modelled data remain independent.

Following Willcock *et al*. (2023), bespoke catchments were generated with each measurement location as catchment outlet, using the 25-m EU-DEM v1.1. (EEA 2016). We allowed for a maximum deviation distance of 100m from the EU-Hydro River Network Database (Copernicus 2020). Locations with larger distances from the river network were omitted. Larger deviation distances could result in erroneous discharge (e.g., from the main river instead of a tributary). Annual discharge was extracted as the sum of all within catchment yields per cell. The total number of validation points after these omissions was 2,252.

We compared the total modelled amount of N (fertiliser, manure, atmospheric and human waste) and P (fertiliser, manure and human waste) summed within the catchments to the amount of measured nutrient estimated to pass through the catchment outlet. For the correct unit of amount, model output accuracy was assessed with the Inverse of Deviance ($D^{\downarrow}$), following Willcock *et al*. (2019): the inverse of the mean individual data point deviances against the 1-1 of equal relative value, and related to Mean Absolute Error. A $D^{\downarrow}$ above 0.7 is seen as significant. Moreover, we conducted rank-order comparisons using Spearman's *Rho* to test for the right order of outcomes. Validation was conducted through bootstraps of 100 datapoints each, for 10,000 runs.

### Results and discussion

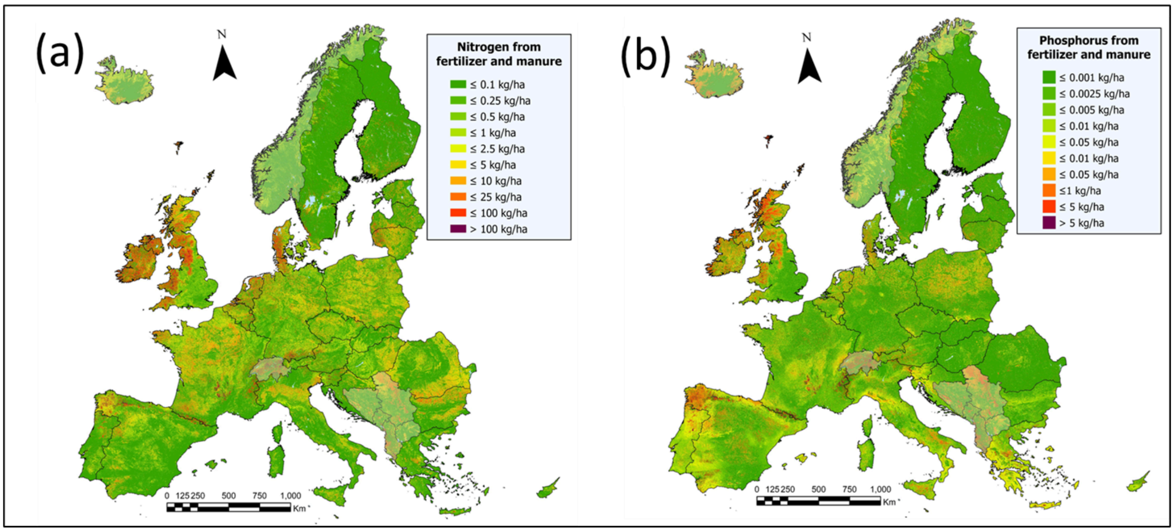


***Figure 2.2:*** *Europe-wide estimate of nutrient losses to the stream from fertiliser and manure additions generated with the InVEST Nutrient Delivery Ratio model for (a) Nitrogen and (b) Phosphorus. Certain non-EU countries are shaded, reflecting that these had no data allowing distinction of crops and estimates were based on median (former Yugoslavia, Switzerland) or Swedish load values (Norway, Iceland).*

Figure 2.2 presents Europe-wide estimates of nutrient losses from fertilisers and manure to waterways with the InVEST Nutrient Delivery Ratio model. Nitrogen is a combination of $NO_x$ and $NH_x$, while P is applied as $P_2O_5$ but modelled as P (Figure 2.3b). Nutrient losses are large from grasslands with a high stocking rate, such as in The Netherlands, Ireland, western Denmark and parts of France and Germany. Other areas that are subject to high nutrient export are grasslands in which a rugged topography is combined with high rainfall, such as in the northern UK, north-western Spain and in Switzerland. Low amounts of nutrient loss are found in regions without widespread intensive agriculture such as in Scandinavia (except Denmark) and southern European regions. In substantial parts of southern Europe, the low loads and nutrient export are related to low to intermediate fertiliser levels for permanent crops, such as olive and fruit trees.

The modelled median for N among catchments was similar to the median of the validation set, 3.57 vs. 2.03 kg per hectare of catchment, as was the comparison of modelled and measured medians for P, 0.07 vs. 0.11 kg/ha P. However, the range of predictions of N was larger than that of the validation set (Table 2.2), whereas the opposite was the case for P – seemingly caused by a set of high validation outliers (Figure 2.3).

| | **Average value modelled data** (5% -95% percentile) | **Average value validation data** (5% -95% percentile) | **Inverse deviance, $D^{\downarrow}$** | **Spearman Rho** |
|---|---|---|---|---|
| **Nitrogen** from fertiliser, manure, atmospheric deposition and human waste | 3.47 kg/ha (0.37-23.6) | 2.03 kg/ha (0.03-12.5) | 0.82 (0.02) | 0.65 (0.06)*** |
| **Phosphorus** from fertiliser and manure | 0.07 kg/ha (0.01-0.62) | 0.11 kg/ha (0.01-0.97) | 0.86 (0.02) | 0.57 (0.09)*** |

***Table 2.2:*** *Validation results of InVEST NDR modelled outputs against EIONET measurements as kg per hectare catchment. Median (±std) among 10,000 bootstrap runs of 100 points each of total n of 2252 across Europe; *** P < 0.001.*

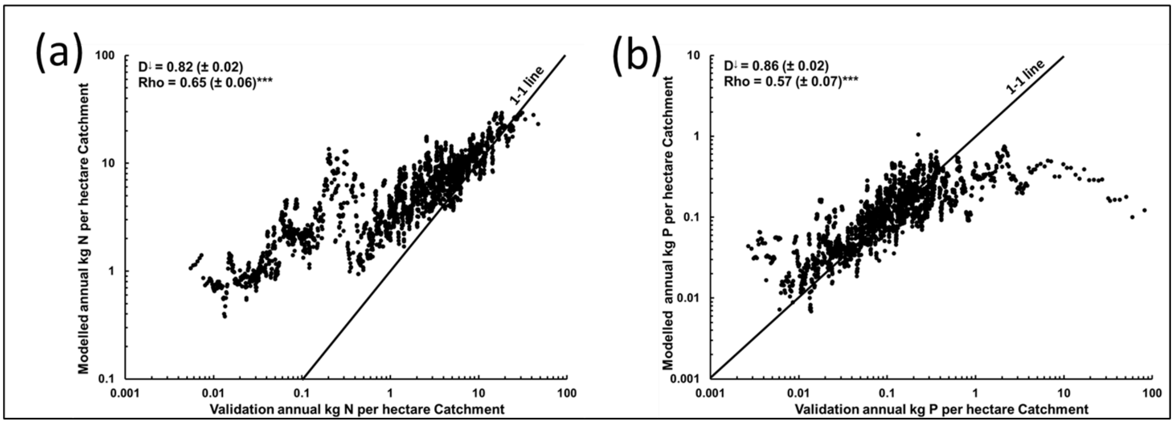

***Figure 2.3:*** *The modelled annual kg per hectare catchment passing through validation locations compared to the validation data as annual kg per hectare. For (a) Nitrogen and (b) Phosphorus. n = 2,252. For depiction purposes, the figures are smoothed along the x-axis with a running average of 10-points.*

Figure 2.3 depicts the validation results. The validation of N is good, with a high inverse deviance (D↓=0.82): showing a close proximity to the 1:1 line. The lower values are overestimated, which is emphasised by the log-log plots. This relates mainly to Scandinavian catchments with low measured concentrations being overestimated by the models. The Spearman rank correlation (Rho) is highly significant ($P \approx 2 \times 10^{-8}$), showing that the order of N export among catchments is well predicted. For P validation is also good. Except for a set of high values, the proximity to the 1-1 line is high (D↓ =0.86). The Spearman rank correlation (Rho) is highly significant ($P \approx 4.5 \times 10^{-10}$).

These results suggest generally good performance of the InVEST NDR model, despite it being a rather simple process-based model. The model does not include in-steam processes, which may explain the poorer match to validation results at low values for N. If, as is likely, there are in-stream denitrification processes, this would result in an extra, unmodelled removal of N, and this resulting decrease in measured in-stream N would be more apparent at low values of N, as seen in Figure 2.3a. The large discrepancy between the model and measured P values at high levels of measured P may result from the simple way in which P export is modelled by InVEST. Phosphorus is often bound to soil particles, which means soil erosion is a major export source. As a result, extra processes such as bank erosion may cause higher inputs than modelled especially in larger catchments.

These possibilities will be addressed in further work in which we will aim to identify and interpret the sources of deviances between modelled and validation values by correlating these deviance values with particular explanatory variables. A major aspect will be to investigate the role of AES, specifically in teams of the ability of the area under AES in catchments to explain variation in deviance. Related processes that might help explain deviances between modelled and measured values are farming practices, livestock density, and cropping systems. Furthermore, the spatial distribution of different vegetation types and type of crops with high or low fertiliser rates could affect the deviances by not accurately catching combined retention capacity. As described above, in-stream processes could also lead to overestimation of modelled values. These can be investigated by looking correlations between deviances and amount and seasonality of discharge, the slope and sinuosity of streams affecting sedimentation, as well as within stream chemical conditions affected by e.g., the pH, amount of reactive dissolved oxygen and presence of substantial amounts of other nutrients and organic carbon.

## References

d' Andrimont, R., et al. (2021). From parcel to continental scale–A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. Remote Sensing of Environment 266, 112708.

Ballabio, C., et al. (2016). Mapping topsoil physical properties at European scale using the LUCAS database. Geoderma 261, 110¬123.

Copernicus (2016). EU-DEM (raster) - version 1.1. [Data-Set] https://sdi.eea.europa.eu/catalogue/srv/api/records/3473589f-0854-4601-919e-2e7dd172ff50

Copernicus (2020). Hydro River Network Database 2006-2012, Europe. [Data-Set] https://doi.org/10.2909/393359a7-7ebd-4a52-80ac-1a18d5f3db9c

Einarsson, R., et al. (2021). Crop production and nitrogen use in European cropland and grassland 1961–2019. Scientific Data 8, 288. [Data-Set]

European Environment Agency (2020). Concentrations of heavy metals and nutrients in agricultural soils. [Data-Set] https://www.eea.europa.eu/en/datahub/datahubitem-view/edbbd466-b845-4e4c-acf9-905ec5e28766?activeAccordion=1083478%2C1083479

European Environment Agency (2022). Waterbase - Water Quality ICM. [Data-Set] https://sdi.eea.europa.eu/data/bdeadea2-cfaf-4724-b002-816d71c7e361

European Space Agency (2020). ESA WorldCover 2020. [Data-set]. https://worldcover2020.esa.int/downloader

FAO (2015). International Boundaries Polygons Level 0 – GAUL. [Data-Set] https://datacore-gn.unepgrid.ch/geonetwork/srv/api/records/e560f98a-f6e4-41a1-bf6b-e0c99fe75426

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.

Hooftman, D.A.P., et al. (2023). A model of sediment retention by vegetation for Great Britain: new methodologies & validation. BioRxiv, 2023-08.

Lavorel, S., et al. (2022). Templates for multifunctional landscape design. Landscape Ecology 37, 913–934.

Lehner, B., et al. (2008). New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union 89, 93¬–94.

Ludemann, C.I., et al. (2022). Global data on fertilizer use by crop and by country. Scientific data 9, 501. [Data-Set]

New, M., et al. (2002). A high-resolution data set of surface climate over global land areas. Climate Research 21, 1¬25.

MODIS (2023a). MOD13Q1 v061: MODIS/Terra Vegetation Indices 16-Day L3 Global 250 m SIN Grid. [Data-Set] https://lpdaac.usgs.gov/products/mod13q1v061/

MODIS (2023b). MOD15A2H v061: MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500 m SIN Grid. [Data-Set] https://lpdaac.usgs.gov/products/mod15a2hv061/

MODIS (2023c). MOD16A2 v061: MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500 m SIN Grid. [Data-Set] https://lpdaac.usgs.gov/products/mod16a2v061/

Redhead, J.W., et al. (2018). National scale evaluation of the InVEST nutrient retention model in the United Kingdom. Science of the Total Environment 610, 666–677.

Robinson, T.P., et al. (2014). Mapping the global distribution of livestock. PLoS one 9, e96084.
Ross, C.W., et al. (2018). Global hydrologic soil groups (HYSOGs250m) for curve number-based runoff modeling. ORNL DAAC, Oak Ridge, Tennessee, USA.
Ruesch, A., & Gibbs, H. (2008). New global biomass carbon map for the year 2000 based on IPCC tier-1 methodology. Carbon Dioxide Information Analysis Center. Oak Ridge National Laboratory, USA.
USDA (2004). Hydrologic Soil-Cover Complexes. In Part 630 Hydrology National Engineering Handbook, Chapter 9. USDA. https://directives.sc.egov.usda.gov/17758.wba
Willcock, S., et al. (2019). A continental-scale validation of ecosystem service models. Ecosystems 22, 1902–1917.
Willcock, S., et al. (2023). Model ensembles of ecosystem services fill global certainty and capacity gaps. Science Advances 9, eadf5492.
Zawadzka, J., et al. (2019). Ecosystem services from combined natural and engineered water and wastewater treatment systems: Going beyond water quality enhancement. Ecological Engineering 142, 100006.

# 3. Improving the Soil Carbon model

The BESTMAP soil carbon model relies on a relatively small number of studies demonstrating the impact of particular AES on soil carbon. To test if this can be improved, we later carried out an extensive review of management practices in agricultural land from European field sites in Cool Temperate and Mediterranean climate zones. We collected and collated data to quantify soil carbon stock changes over a range of sample depths from 0 – 150 cm and time periods between 1 and 40 years. The methodology for the analysis is described in detail in a forthcoming publication (email c.willoughby@leeds.ac.uk for more information regarding the methods). Management practices included cover cropping, organic farming, planting trees on arable land and incorporation of ley periods in arable rotations.

## Carbon accumulation and stock change results

Results showed that soil carbon stock increases were enacted through practices which introduced and maintained a diverse community of living roots in agricultural land. Specifically, included the conversion of arable land to forestry, keeping soil covered through cover crops, including leys and fallow periods in crop rotations and maintaining grassland through grazing livestock (Table 3.1). We found that conversion to organic farm management resulted in losses of soil carbon (Table 3.1).

| Management Practice | Mean Carbon Accumulation Rate (annual t C ha $^{-1}$) [1st, 3rd Quartile] |
| --- | --- |

| | |
|---|---|
| Forestry Conversion | 0.06 [0.00, 0.31] |
| Cover Crops | 0.05 [0.00, 0.23] |
| Leys and Fallow Periods | 0.21 [0.02, 0.34] |
| Maintaining Grassland | 0.23 [0.00, 0.34] |
| Organic Farming | -0.29 [-0.55, 0.08] |

***Table 3.1:*** *A summary of carbon accumulation rate according to selected management practices.*

The results showed that rates of accumulation were variable between the management factors incorporated in the analysis. The introduction of leys into arable rotations and use of fallow periods on cropland had consistently positive carbon accumulation rates (Table 3.1). For the majority of the other management groups, results ranged between no change in carbon accumulation up to 0.34 t C annually for activities relating to leys and grassland maintenance through incorporation of grazing livestock (Table 3.1). Even though results from conversion from conventional agriculture to organic management were generally negative (Table 3.1), not all organic conversion studies reported a negative transition, with a wider variability in results identified as study length increased and a lack of long term studies to draw from (Figure 3.1).

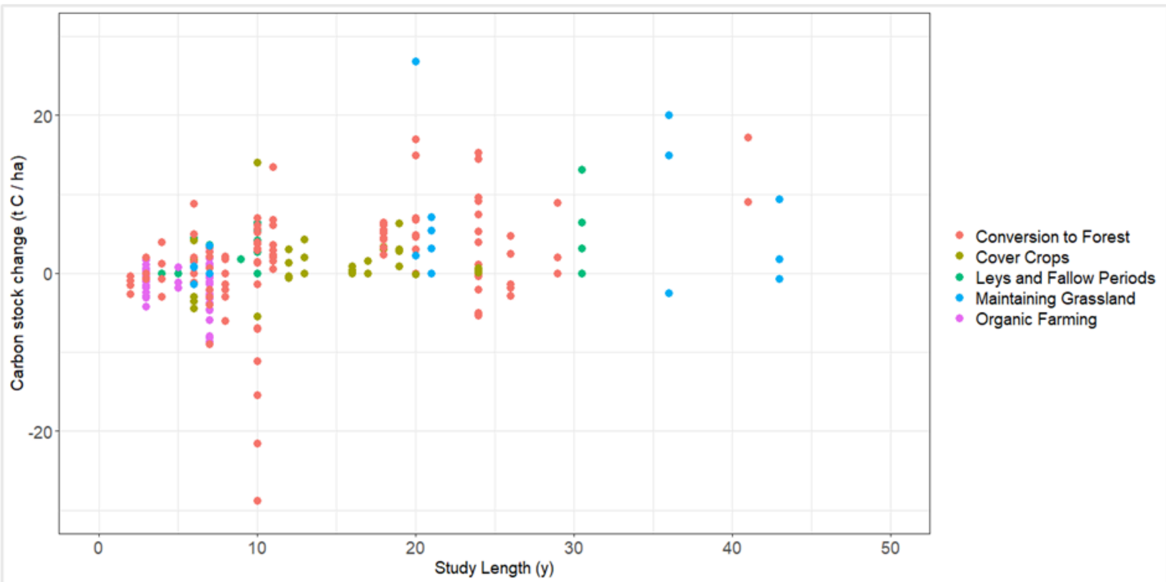

***Figure 3.1:*** *carbon stock change and length of study from the analysis, coloured according to management type.*

Results showed that stock changes through time varied with each management practice (Figure 3.1). Stock changes through conversion to forest varied in studies up to 12 years in duration, but after this point there was a generally positive trend – we found that in sites with a duration of 40 years post forestry conversion, up to 18 t C ha $^{-1}$ was added to the soil. The introduction of leys and fallow periods was found to increase with study duration, with a maximum C stock change of 13 t ha $^{-1}$ in a study with a 31-year duration. Activities which maintained grassland varied in stock change through time, with a peak of 27 t C ha $^{-1}$ found in sites 20 years post grassland-maintaining activities. Our results showed a potential decrease in stock change as grassland maintenance passed beyond 20 years in duration (Figure 3.1). Cover cropping and grassland maintenance had the most consistently positive stock changes, and cover cropping stock changes remained consistent through time with no large increases in soil C found as study length increased. The peak soil C stock change found underlying cover cropping was 14 t ha $^{-1}$ in a study with a 10-year duration.

We found that the capacity for management changes to elicit C stock changes was limited by sample depth (Figure 3.2). Stock changes were more variable in the topsoil, and results showed consistent decreases in magnitude as sample depth increased. Few studies sampled below 50 cm, and those that were included in our analysis were limited to forestry conversion and cover cropping studies (Figure 3.2).

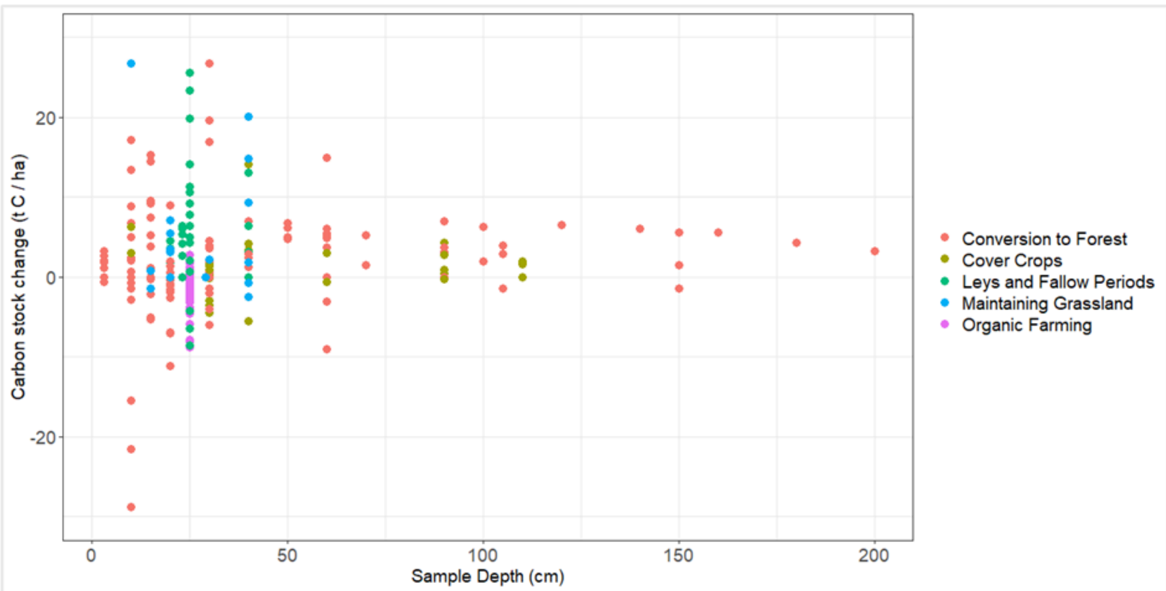


***Figure 3.2:*** *carbon stock change and sample depth from the analysis, coloured according to management type.*

### Research Needs

Our work has identified several gaps in the current body of literature that should be addressed by future research, which may be summarised into five key areas of concern (Figure 3.3). We divided issues into those which should be addressed at an international scale and those requiring further field-scale research. We consider the resilience of soils both as a linkage between the scales and furthermore as the foundation for future land management, given the likelihood of increasingly extreme weather events due to climate change. In the future, assessing soil resilience to challenging conditions alongside its capability to store carbon will help to ensure the sustainability of land management practices, accounting for the need for soil to fulfil multiple roles and ecosystem services. We recommend that soil sampling campaigns include physical indicators alongside carbon content to enable stock changes to be better quantified across a wider area.

Our analysis showed that much of our knowledge on soil carbon stock change is limited to the top 50 cm of the soil because this is where many of our previous studies have been confined. The same is true of national and international soil sampling campaigns, many of which do not collect data below 30 cm of soil. This is a huge gap in our understanding of interactions between management practices and soil carbon stocks. The clear solution to this is to ensure that future studies include subsoil analysis as standard in their sampling, and for any national-level soil inventory schemes to include subsoil analysis in their projects. This would ensure that the possibility of soil carbon redistribution through the soil profile is accounted for in carbon stock change inventories.

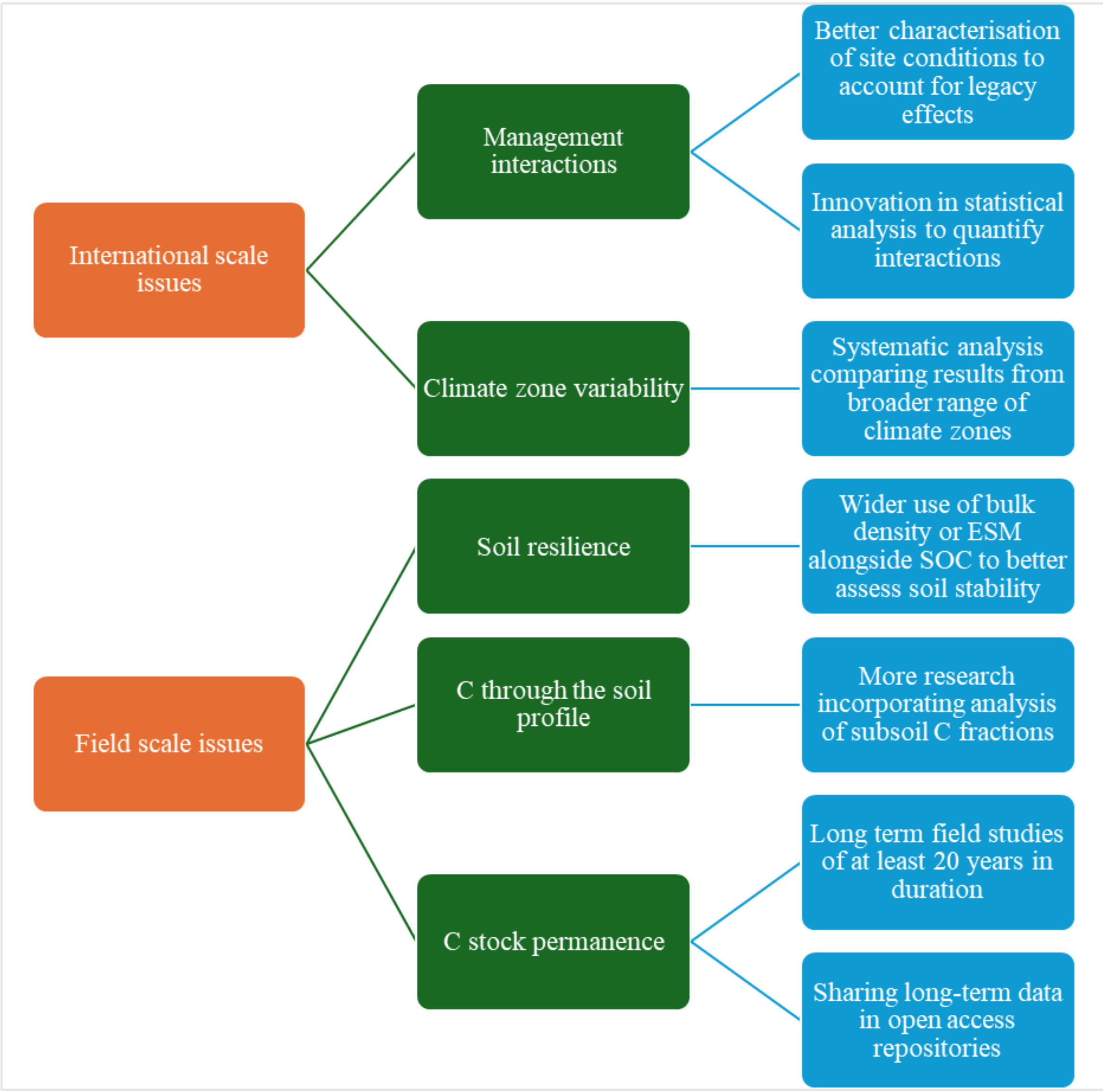


***Figure 3.3:*** *A summary of issue scales (orange), research needs (green) and proposed future solutions (blue).*

Understanding soil carbon distribution through the profile would enable researchers to make evidence-driven management recommendations, however there is a need for a development in our understanding of additionality of management practices and how they may impact carbon stocks in the field. The studies included in this analysis considered a maximum of three management factors at any given site, but we understand that in a commercial setting there are many combinations of management practices, which will be influenced by specific climate, physical or economic factors. There is a need for more research and development of the current base of field sites across Europe to account for legacy effects of previous management and potentially the need for management to change across field sites to account for more realistic scenarios in a commercial setting. In tandem with this, there is a need to work with statisticians and data analysts to develop and validate models and forms of analysis to investigate the effects of management interactions on underlying soil carbon stocks.

Finally, there is already an awareness among researchers that carbon stock changes do not equate to permanent storage or offsetting. However, there is a lack of clarity of terms, with carbon stocks referred to as "stocks", "storage", "accumulation" and "sequestration" while referring to the same type of data. This is naturally confusing and creates unneeded uncertainty in the understanding of soil carbon stock dynamics, as permanence of the C that has been added to the soil may be poorly quantified. The solutions to this issue may already exist in the form of long-term data from field experiments, which can be used as long-term records of soil carbon changes underlying contrasting management strategies across decades or even hundreds of years. It is therefore important to encourage the sharing of information from these long-term experiments between institutions and countries to ensure that researchers are able to access the most relevant information possible with regard to long term soil carbon storage.

## References

Aubrion, G., Fontaine, A., Kerveillant, P., Beaudoin, N., Constantin, J., Mary, B., Laurent, F. (2010). Effects of catch crops, no till and reduced nitrogen fertilization on nitrogen leaching and balance in three long-term experiments. Agriculture, Ecosystems & Environment. https://doi.org/10.1016/j.agee.2009.10.005

Beckert, M. R., Smith, P., Lilly, A., Chapman, S. J. (2016). Soil and tree biomass carbon sequestration potential of silvopastoral and woodland-pasture systems in North East Scotland. Agroforest Systems. https://doi.org/10.1007/s10457-015-9860-4

Breil, N. L., Lamaze, T., Bustillo, V., Marcato-Romain, C.-E., Coudert, B., Queguiner, S., Jarosz-Pelle, N. (2023). Combined impact of no-tillage and cover crops on soil carbon stocks and fluxes in maize crops. https://doi.org/10.1016/j.still.2023.105782

Cardinael, R., Chevallier, T., Barthès, B. G., Saby, N., Parent, T., Dupraz, C., Bernoux, M., Chenu, C. (2015). Impact of alley cropping agroforestry on stocks, forms and spatial distribution of soil organic carbon — A case study in a Mediterranean context. https://doi.org/10.1016/j.geoderma.2015.06.015

Cardinael, R., Chevallier, T., Cambou, A., Béral, C., Barthès, B. G., Dupraz, C., Durand, C., Kouakoua, E., Chenu, C. (2016). Increased soil organic carbon stocks under agroforestry: A survey of six different sites in France. Agriculture, Ecosystems and Environment. https://doi.org/10.1016/j.agee.2016.12.011

Constantin, J., Mary, B., Laurent, F., Aubrion, G., Fontaine, A., Kerveillant, P., Beaudoin, N. (2010). Effects of catch crops, no till and reduced nitrogen fertilization on nitrogen leaching and balance in three long-term experiments. Agriculture, Ecosystems & Environment. https://doi.org/10.1016/j.agee.2009.10.005

Del Galdo, I., Six, J., Peressotti, A., Cotrufo, M. F. (2003). Assessing the impact of land-use change on soil C sequestration in agricultural soils by means of organic matter fractionation and stable C isotopes. Global Change Biology. https://doi.org/10.1046/j.1365-2486.2003.00657.x

Fornara, D. A. (2018). Land use change and soil carbon pools: evidence from a long-term silvopastoral experiment. Agroforestry Systems. https://doi.org/10.1007/s10457-017-0124-3

Glisczynski, F. V., Pude, R., Amelung, W., Sandhage-Hofmann, A. (2016). Biochar-compost substrates in short-rotation coppice: Effects on soil and trees in a three-year field experiment. Journal of Plant Nutrition and Soil Science. https://doi.org/10.1002/jpln.201500545

Heyburn, J., McKenzie, P., Crawley, M. J., Fornara, D. A. (2017). Effects of grassland management on plant C: N: P stoichiometry: implications for soil element cycling and storage. https://doi.org/10.1002/ecs2.1963

Hu, T., Sorensen, P., Olesen, J. E. (2018). Soil carbon varies between different organic and conventional management schemes in arable agriculture. European Journal of Agronomy. https://doi.org/10.1016/j.eja.2018.01.010

Johnston, A. E., Poulton, P. R., Coleman, K., Macdonald, A. J., White, R. P. (2017). Changes in soil organic matter over 70 years in continuous arable and ley–arable rotations on a sandy loam soil in England. European Journal of Soil Science. https://doi.org/10.1111/ejss.12415

Medina-Roldán, E., Paz-Ferreiro, J., Bardgett, R. D. (2012). Grazing exclusion affects soil and plant communities but has no impact on soil carbon storage in an upland grassland. Agriculture, Ecosystems & Environment. https://doi.org/10.1016/j.agee.2011.12.012

Medinski, T. V., Freese, D., Böhm, C., Slazak, A. (2014). Soil carbon fractions in short rotation poplar and black locust coppices, Germany. Agroforest Systems. https://doi.org/10.1007/s10457-014-9709-2

Peregrina, F. (2019). Soil carbon content and its stratification at the medium term (5 and 8 years) in a semiarid vineyard with cover crops. Spanish Journal of Soil Science, 9(2), 10.3232/SJSS.2019.V9.N2.01.

Plaza-Bonilla, D., Nolot, J.-M., Passot, S., Raffaillac, D., Justes, E. (2016). Grain legume-based rotations managed under conventional tillage need cover crops to mitigate soil organic matter losses. Soil & Tillage Research. https://doi.org/10.1016/j.still.2015.09.021

Poulton, P., Johnston, J., MacDonald, A., White, R., Powlson, D. (2018). Major limitations to achieving "4 per 1000" increases in soil organic carbon stock in temperate regions:

evidence from Long term experiments at Rothamsted research, United Kingdom. https://doi.org/10.111/gcb.14066

Rasmussen, J., Eriksen, J., Hansen, E. M., Christensen, B. T. (2008). Carbon sequestration and residual effect of differently aged grass leys. Proceedings of NJF Seminar 407.

Schulz, F., Gattinger, A., Brock, C., Leithold, G. (2017). Organic farming with livestock raising vs. stockless farming - development of soil organic matter stocks and cash crop yields. Scientific Track “Innovative Research for Organic Agriculture 3.0” 19th Organic World Congress, New Delhi, India, November 9-11, 2017 Organized by ISOFAR, NCOF and TIPI.

Smith, S. W., Vandenberghe, C., Hastings, A., Johnson, D., Pakeman, R. J., Van Der Wal, R., Woodin, S. J. (2013). Optimizing Carbon Storage Within a Spatially Heterogeneous Upland Grassland Through Sheep Grazing Management. DOI: 10.1007/s10021-013-9731-7.

Upson, M.A., Burgess, P.J., Morison, J.I.L. (2016). Soil carbon changes after establishing woodland and agroforestry trees in a grazed pasture. Geoderma. http://dx.doi.org/10.1016/j.geoderma.2016.07.002

Viaud, Kunnemann. (2021). Additional soil organic carbon stocks in hedgerows in crop-livestock areas. Agriculture, Ecosystems & Environment. https://doi.org/10.1016/j.agee.2020.107174

## 4. Remote Sensing of Soil Organic Carbon

The BESTMAP project made extensive use of Top Soil Carbon in the ABM modelling (see Deliverable 4.1). Given the importance of that data, and the coarse nature of the existing SOC remote sensing, a pilot was conducted to explore how we can improve this estimate as part of Task 5.3.

Soil plays a key role in the global carbon cycle. It is estimated that the top 30cm of soil store 700 Pg soil organic carbon (SOC) globally - this is almost equivalent to the quantity of carbon stored in the atmosphere (870 Pg C in 2018) (Kopittke et al., 2019). As a result, changes in SOC stock can affect atmospheric $CO_2$ concentration (Kätterer et al., 2012). Increasing SOC stocks in agricultural land is considered an attractive negative emission technology which yields benefits for soil quality and fertility and does not require further land-use conversions (Paustian et al., 2019). In fact, global croplands cover 1,410 million ha and are estimated to stock 83 Pg C in the top 30cm of soil (Padarian et al., 2022). Although reliable estimates remain challenging, global croplands could stock an additional 29 Pg C to 65 Pg C. This could offset three to seven years of global anthropogenic emissions (estimated to be at 37.4 Pg $CO_2$ eq. $yr^{-1}$) (Friedlingstein et al., 2022; Padarian et al., 2022).

In light of climate change and global population increase, there is a growing need to monitor SOC stocks in agricultural land to ensure that SOC is not being lost, to track SOC change after the adoption of more sustainable soil management practices and, lastly, to comply with policy frameworks. Traditional soil surveys for direct SOC content and stock measurement are expensive and time-consuming. Hence why, in recent years, remote sensing has gained popularity in complementing soil sampling campaigns for rapid, large-scale and spatially continuous SOC content prediction. Remote sensing of SOC is an indirect quantification technique that consists in training statistical or machine learning models that predict the carbon content of agricultural soil from its satellite reflectance.

SOC does not have any specific signatures across the satellite spectra. Instead, higher SOC contents decrease the overall reflectance of soil. However, two other factors can affect soil reflectance in the same way: crop residue cover and soil moisture. Additionally, clouds, variable sun illumination angles and the adoption of different satellite image pre-processing algorithms can all introduce artifacts to the satellite reflectance, potentially leading to skewed SOC predictions (see Vaudour et al., 2022). Due to these factors, remote sensing-based SOC content predictions are often very inaccurate, especially over large geographical scales (Tziolas et al., 2021).

### Optimal bare soil satellite extraction over European agricultural land

This work tested existing and new approaches of obtaining bare soil satellite reflectance from ESA's Sentinel-2 mission. The aim was to identify which combination of filtering and data aggregation techniques leads to a closer match to laboratory soil spectra, which should lead to improved SOC content predictions.

The first step involved investigating the LUCAS 2018 picture database. LUCAS is an in-situ data collection campaign run every 3 years by Eurostat aimed at collecting land cover information and standardised statistics across the EU. In particular, the LUCAS TOPSOIL module focuses on gathering data on soil properties (Fernández-Ugalde et al., 2022). LUCAS 2018 points with the following characteristics were selected:

- soil sample collected as part of LUCAS TOPSOIL 2018
- belonging to agricultural land use (cropland and fallow land)
- bare (or crop residue-covered) at the time of survey in 2018
- located more than 20m away from the field boundary (to avoid Sentinel-2 pixel boundary mixing).

This resulted in 408 LUCAS 2018 TOPSOIL points (see Figure 4.1), of which 272 were covered in no or very low amounts of crop residue cover (hereby classified as "bare").

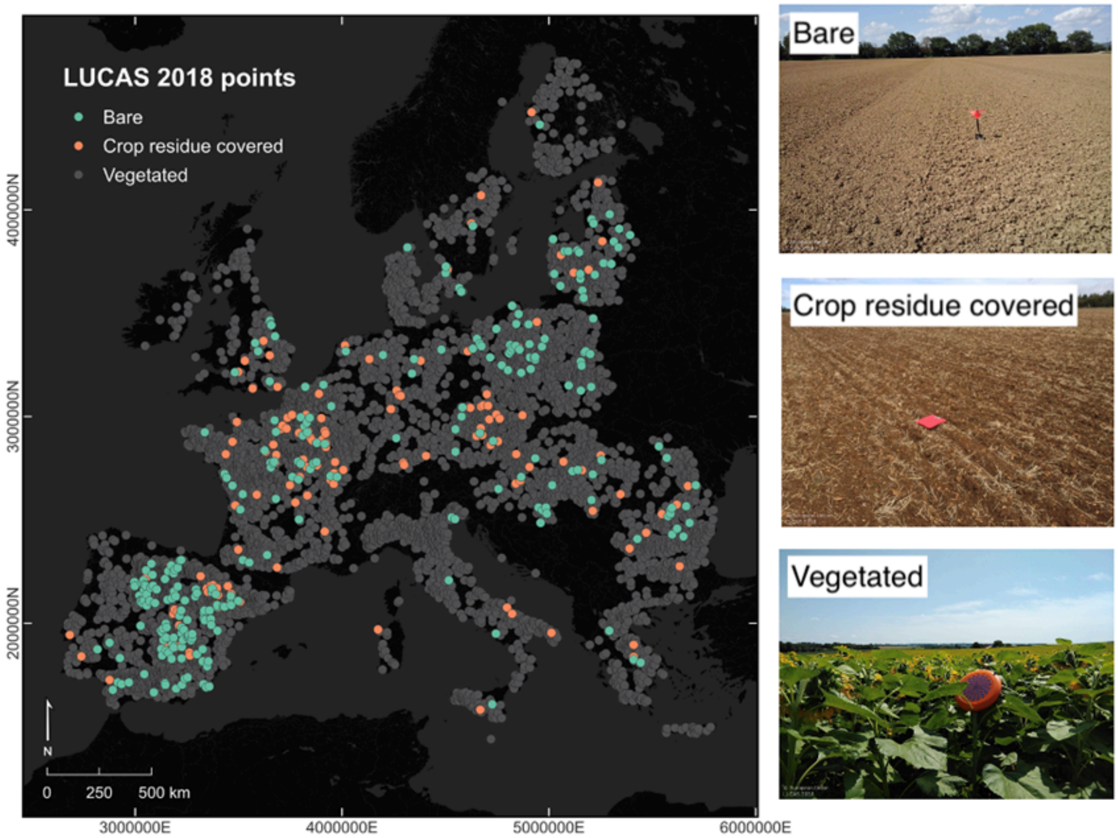


***Figure 4.1:*** *Location of the bare, crop residue-covered and vegetated points in the LUCAS 2018 picture database.*

For each of the 272 bare points, Sentinel-2 surface reflectance imagery taken 2 months before and after the point survey date were selected. In this step, three separate Sentinel-2 datasets were used to test the influence of different image pre-processing algorithms: sen2cor-processed imagery extracted from the Google Earth Engine catalogue (cloud-masked with S2cloudless), FORCE-processed imagery with BRDF correction and FORCE-processed imagery without BRDF correction. The latter two were processed by Mundialis Gmbh. We compare sen2cor and FORCE (Framework for Operational Radiometric Correction for Environmental monitoring) data to assess the effect of different processing algorithms, which might introduce artefacts especially in Shortwave Infrared bands B11, B12, the most important bands for SOC content prediction. A direct comparison of FORCE and sen2cor B11 and B12 data has not been carried out (Skakun et al., 2022). It is to be noted that FORCE allows an additional Bidirectional Reflectance Distribution Function (BRDF) correction feature to sen2cor, which can correct for variable solar illumination angle (Franch et al., 2019; Frantz, 2019). Additionally, FORCE-processed data has not yet been used for SOC content prediction.

LUCAS 2015 Ancillary information such as soil type (Jones et al., 2020) and daily volumetric soil moisture content from the Global Surface Soil Moisture dataset (GSSM1km, Han et al., 2023) were added to each image in each image collection. Then, various satellite band ratio thresholds (NDVI to mask out vegetated pixels and NBR2 to mask out crop residue-covered/wet pixels) were tested (see Figure 4.2). Finally, the resulting spectra acquired on different dates for each point were aggregated in various ways:

- Selecting the R90, believed to represent the driest soil conditions while avoiding cloud interferences (the 90th percentile reflectance value, (Castaldi et al., 2023).
- Computing the mean spectra (Diek et al., 2016; Dvorakova et al., 2023; Rogge et al., 2018; Vaudour et al., 2021)
- Selecting the median spectra (Demattê et al., 2018).
- Selecting the driest date spectra (Vaudour et al., 2021, 2019).
- Computing the Synthetic 0% Soil Moisture. For each band, the reflectance of each image was plotted against its volumetric soil moisture (Figure 4.2). An exponential function was applied based on the nonlinear relationship between soil reflectance and soil moisture observed in the laboratory. The exponential function was obtained following the approach by Lobell and Asner (2002), who developed a model of soil reflectance under different moisture conditions for US soil samples. For each point, if the dry reflectance predicted by the function fell within the minimum and maximum range of reflectance for that band and for that soil type according to the LUCAS 2015 spectroscopy database, then the y intercept was used as synthetically dry reflectance in that band for that point. If it fell outside the range, the R90 reflectance was selected instead.

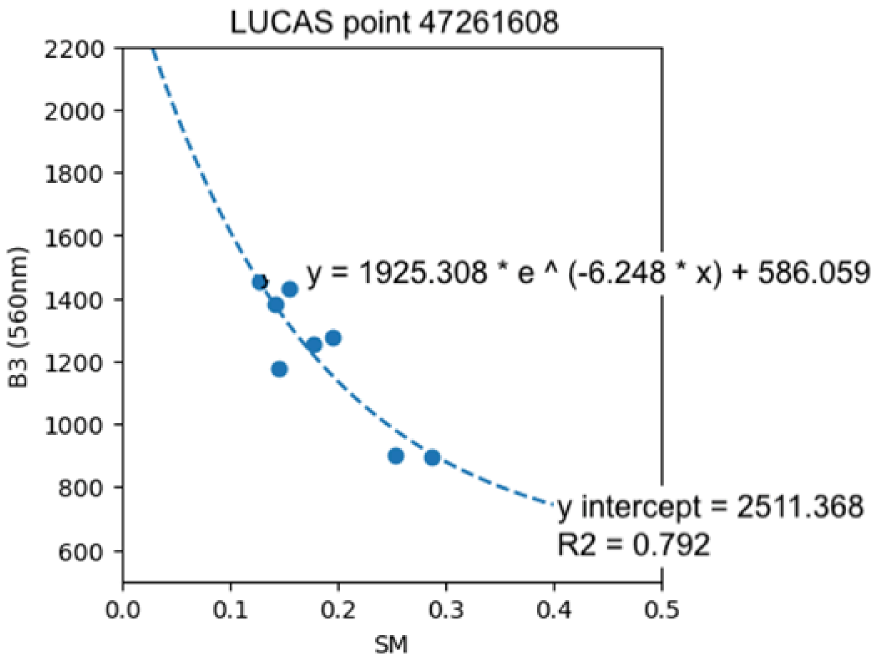


***Figure 4.2:*** *How the synthetically dry reflectance was determined. Example for Band 3 (560 nm) for one LUCAS point.*

The LUCAS 2015 TOPSOIL dataset was filtered to include points that were also surveyed in LUCAS TOPSOIL 2018, and all the points that witnessed a <10% SOC content change between the 2015 and 2018 survey dates were considered. For each of these points, the corresponding LUCAS 2015 laboratory soil reflectance spectrum was convolved to Sentinel-2 reflectance according to the Sentinel-2 spectral response function. We assume this convolved spectra to reflect "ideal" bare soil conditions because, in the lab, soil samples are dried to 0% soil moisture and are not covered by vegetation or crop residue cover (Nocita et al., 2015). Lastly, several metrics (whole-spectra MAE, ANOVA, Tukey's test and band-by-band Pearson's r) were used to determine the match of each aggregated satellite spectra to the laboratory spectra. The methodology is summarised in Figure 4.3.

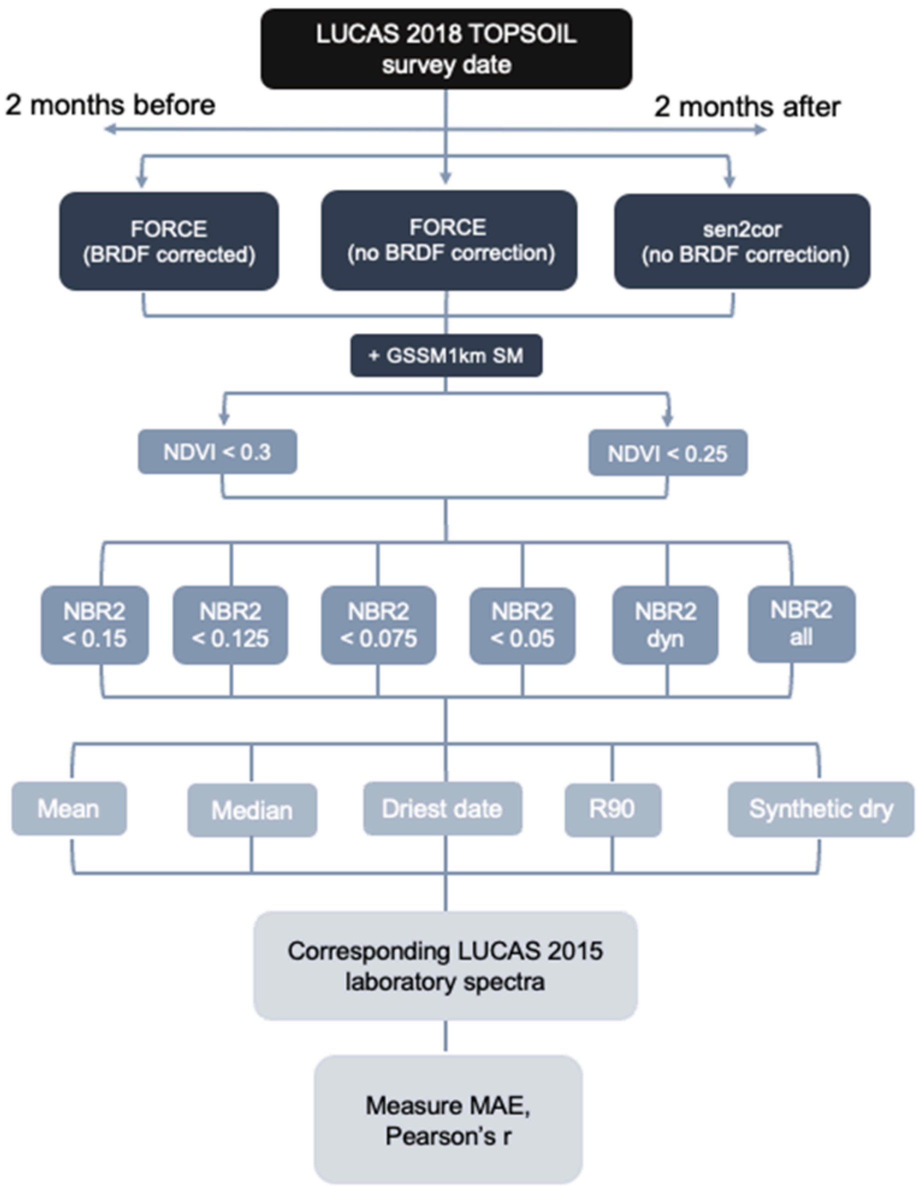


***Figure 4.3:*** *Optimal bare soil satellite reflectance extraction over selected LUCAS points across Europe. The dynamic NBR2 consists in taking the 85$^{th}$ quantile of NBR2 for spread for Acrisols, Calcisols, Leptosols, Regosols and Vertisols in the LUCAS 2015 spectral*

*database, who have relatively high bare soil NBR2 (>0.05), while leaving NBR2 < 0.05 for all other soil types.*

Initial results indicate that, in terms of MAE over the whole spectra, sen2cor data provided a closer match to laboratory spectra than FORCE data (see Figure 4.4). ANOVA and Tukey's testing suggest that, overall, synthetically dry and R90 spectra aggregation techniques resulted in a closer match to laboratory spectra than Driest date, Median and Mean approaches.

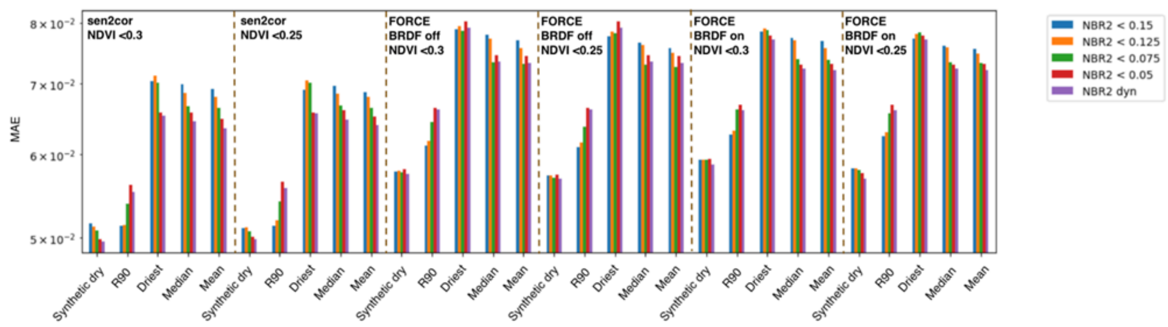


***Figure 4.4:*** *Summary of whole-spectra MAE for each dataset, NDVI and NBR2 threshold tested in this study.*

Band-by-band Pearson's r paints a different picture. All bands show strong (r>0.5) Pearson correlation, except for the NDVI and NBR2 indices of the synthetically dry spectra. However, FORCE data (both BRDF and non-BRDF corrected) returned slightly stronger band-by-band Pearson's correlation coefficients than sen2cor data, especially across bands B2 (490 nm), B3 (560 nm), B8 (842 nm), B11 (1610 nm) and B12 (2190 nm), which also happen to be the most important for SOC content prediction (see Table 4.1, Table 4.2). There doesn't seem to be a major difference between FORCE BRDF corrected and non-BRDF corrected data in terms of band-by-band Pearson's r.

| Method | 490 | 560 | 665 | 705 | 740 | 783 | 842 | 865 | 1610 | 2190 | ndvi | nbr2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SYNTHETIC | 0.577237 | 0.607147 | 0.647500 | 0.647095 | 0.666019 | 0.645143 | 0.610070 | 0.665098 | 0.586586 | 0.488324 | 0.245341 | 0.328430 |
| R90 | 0.749898 | 0.726238 | 0.723781 | 0.726913 | 0.715921 | 0.704382 | 0.680069 | 0.686529 | 0.713214 | 0.624723 | 0.660170 | 0.652217 |
| DRIEST | 0.716055 | 0.711423 | 0.721980 | 0.719245 | 0.702214 | 0.684492 | 0.664362 | 0.661379 | 0.675616 | 0.575196 | 0.616301 | 0.570766 |
| MEDIAN | 0.736714 | 0.715649 | 0.719557 | 0.726414 | 0.720031 | 0.708837 | 0.684193 | 0.686611 | 0.711906 | 0.626143 | 0.684429 | 0.683493 |
| MEAN | 0.757429 | 0.737598 | 0.738629 | 0.743986 | 0.736599 | 0.724780 | 0.701608 | 0.703992 | 0.720289 | 0.625390 | 0.704269 | 0.713657 |

***Table 4.1:*** *Band-by-band Pearson's r of sen2cor data with NDVI < 0.25 and NBR2 < 0.05.*

| Method | 490 | 560 | 665 | 705 | 740 | 783 | 842 | 865 | 1610 | 2190 | ndvi | nbr2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SYNTHETIC | 0.667054 | 0.675886 | 0.636950 | 0.682084 | 0.691597 | 0.687311 | 0.685363 | 0.692262 | 0.720352 | 0.660945 | 0.288725 | 0.153570 |
| R90 | 0.819350 | 0.781036 | 0.746204 | 0.741792 | 0.735607 | 0.727989 | 0.722538 | 0.726198 | 0.774223 | 0.675062 | 0.678474 | 0.548378 |
| DRIEST | 0.763427 | 0.717187 | 0.683279 | 0.689747 | 0.683754 | 0.678072 | 0.668675 | 0.669470 | 0.723952 | 0.604648 | 0.597949 | 0.510926 |
| MEDIAN | 0.802993 | 0.763656 | 0.726393 | 0.728344 | 0.723845 | 0.718645 | 0.708285 | 0.714469 | 0.769648 | 0.654628 | 0.689280 | 0.603603 |
| MEAN | 0.799866 | 0.759976 | 0.725323 | 0.726147 | 0.720385 | 0.714540 | 0.706185 | 0.709265 | 0.765473 | 0.655641 | 0.680252 | 0.600458 |

***Table 4.2:*** *Band-by-band Pearson's r of FORCE BRDF corrected data with NDVI<0.25 and NBR2<0.05.*

### Future research needs

As part of wider efforts to try and increase the accuracy of satellite SOC predictions, this exploratory research investigated how bare soil satellite reflectance could be improved using some established and some innovative satellite soil reflectance extraction techniques.

The model of soil reflectance with changing soil moisture introduced in this work to compute the synthetically dry spectra was developed using a single soil type from the UK (Calcaric Cambisol from the University of Leeds Farm), but the original method was developed on 4 different US soil types (Lobell and Asner, 2002). After Cambisols, the most common LUCAS 2018 soil types dedicated to agricultural or fallow land use are Luvisols, Calcisols, Leptosols and Fluvisols. Therefore, to get a more representative model of soil reflectance with changing soil moisture for the creation of a synthetically dry spectra over the whole of Europe, future research efforts should be aimed at testing and calibrating the reflectance-soil moisture model over these European soil types. Additionally, this work assumes that the best satellite best spectra are the ones most closely resembling convolved laboratory soil reflectance, because the laboratory spectrum represents ideal (dry and uncontaminated) bare soil conditions. However, some studies suggest that a resemblance to laboratory soil spectra does not always lead to improved SOC content predictions (Castaldi, 2021). For this reason, the next steps in this research will involve using the spectra obtained as part of this work for SOC content prediction. The method resulting in highest SOC prediction accuracy will be considered the best bare soil satellite extraction technique.

Lastly, multispectral satellites provide limited spectral resolution, but recent studies show that the improved spectral resolution of hyperspectral satellites (PRISMA, EnMAP) leads to improved SOC content prediction at scale (see Tziolas et al., 2021). Research efforts will need to focus on algorithms for the handling and processing of reflectance data from hyperspectral satellites for SOC content prediction, as these become operational and start gathering data over large areas. For instance, ESA is planning to launch a hyperspectral satellite (CHIME) as part of the Copernicus program for routine hyperspectral observations in the late 2020s (Nieke and Rast, 2018).

## References

Castaldi, F., 2021. Sentinel-2 and landsat-8 multi-temporal series to estimate topsoil properties on croplands. Remote Sens. 13, 3345.

Castaldi, F., Koparan, M.H., Wetterlind, J., Žydelis, R., Vinci, I., Savaş, A.Ö., Kıvrak, C., Tunçay, T., Volungevičius, J., Obber, S., 2023. Assessing the capability of Sentinel-2 time-series to estimate soil organic carbon and clay content at local scale in croplands. ISPRS J. Photogramm. Remote Sens. 199, 40–60.

Demattê, J.A.M., Fongaro, C.T., Rizzo, R., Safanelli, J.L., 2018. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. Remote Sens. Environ. 212, 161–175.

Diek, S., Schaepman, M.E., De Jong, R., 2016. Creating multi-temporal composites of airborne imaging spectroscopy data in support of digital soil mapping. Remote Sens. 8, 906.

Dvorakova, K., Heiden, U., Pepers, K., Staats, G., van Os, G., van Wesemael, B., 2023. Improving soil organic carbon predictions from a Sentinel–2 soil composite by assessing surface conditions and uncertainties. Geoderma 429, 116128. https://doi.org/10.1016/j.geoderma.2022.116128

Fernández-Ugalde, O., Scarpa, S., Orgiazzi, A., Panagos, P., Van Liedekerke, M., Marechal, A., Jones, A., 2022. LUCAS 2018 Soil Module. Presentation of dataset and results, EUR 31144 EN, Publications Office of the European Union. Luxembourg. Doi 10, 215013.

Franch, B., Vermote, E., Skakun, S., Roger, J.-C., Masek, J., Ju, J., Villaescusa-Nadal, J.L., Santamaria-Artigas, A., 2019. A method for Landsat and Sentinel 2 (HLS) BRDF normalization. Remote Sens. 11, 632.

Frantz, D., 2019. FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. Remote Sens. 11, 1124. https://doi.org/10.3390/rs11091124

Friedlingstein, P., Jones, M.W., O'Sullivan, M., Andrew, R.M., Bakker, D.C.E., Hauck, J., Le Quéré, C., Peters, G.P., Peters, W., Pongratz, J., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Anthoni, P., Bates, N.R., Becker, M., Bellouin, N., Bopp, L., Chau, T.T.T., Chevallier, F., Chini, L.P., Cronin, M., Currie, K.I., Decharme, B., Djeutchouang, L.M., Dou, X., Evans, W., Feely, R.A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Luijkx, I.T., Jain, A., Jones, S.D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Körtzinger, A., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P.C., Melton, J.R., Munro, D.R., Nabel, J.E.M.S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T.M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A.J., Sweeney, C., Tanhua, T., Tans, P.P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G.R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A.J., Willis, D., Wiltshire, A.J., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., 2022. Global Carbon Budget 2021. Earth Syst. Sci. Data 14, 1917–2005. https://doi.org/10.5194/essd-14-1917-2022

Kätterer, T., Bolinder, M.A., Berglund, K., Kirchmann, H., 2012. Strategies for carbon sequestration in agricultural soils in northern Europe. Acta Agric. Scand. Sect. A–Animal Sci. 62, 181–198.

Kopittke, P.M., Menzies, N.W., Wang, P., McKenna, B.A., Lombi, E., 2019. Soil and the intensification of agriculture for global food security. Environ. Int. 132, 105078. https://doi.org/10.1016/j.envint.2019.105078

Lobell, D.B., Asner, G.P., 2002. Moisture Effects on Soil Reflectance. Soil Sci. Soc. Am. J. 66, 722–727. https://doi.org/10.2136/sssaj2002.7220

Nieke, J., Rast, M., 2018. Towards the copernicus hyperspectral imaging mission for the environment (CHIME), in: Igarss 2018-2018 Ieee International Geoscience and Remote Sensing Symposium. IEEE, pp. 157–159.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Ben Dor, E., Brown, D.J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J.A.M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai, H., Soriano-Disla, J.M., Shepherd, K.D., Stenberg, B., Towett, E.K., Vargas, R., Wetterlind, J., 2015. Chapter Four - Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring, in: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 139–159. https://doi.org/10.1016/bs.agron.2015.02.002

Padarian, J., Minasny, B., McBratney, A., Smith, P., 2022. Soil carbon sequestration potential in global croplands. PeerJ 10, e13740. https://doi.org/10.7717/peerj.13740

Paustian, K., Larson, E., Kent, J., Marx, E., Swan, A., 2019. Soil C sequestration as a biological negative emission strategy. Front. Clim. 8.

Rogge, D., Bauer, A., Zeidler, J., Mueller, A., Esch, T., Heiden, U., 2018. Building an exposed soil composite processor (SCMaP) for mapping spatial and temporal characteristics of soils with Landsat imagery (1984–2014). Remote Sens. Environ. 205, 1–17. https://doi.org/10.1016/j.rse.2017.11.004

Skakun, S., Wevers, J., Brockmann, C., Doxani, G., Aleksandrov, M., Batič, M., Frantz, D., Gascon, F., Gómez-Chova, L., Hagolle, O., 2022. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. Remote Sens. Environ. 274, 112990.

Tziolas, N., Tsakiridis, N., Chabrillat, S., Demattê, J.A.M., Ben-Dor, E., Gholizadeh, A., Zalidis, G., van Wesemael, B., 2021. Earth Observation Data-Driven Cropland Soil Monitoring: A Review. Remote Sens. 13, 4439. https://doi.org/10.3390/rs13214439

Vaudour, E., Gholizadeh, A., Castaldi, F., Saberioon, M., Borůvka, L., Urbina-Salazar, D., Fouad, Y., Arrouays, D., Richer-de-Forges, A.C., Biney, J., Wetterlind, J., Van Wesemael, B., 2022. Satellite Imagery to Map Topsoil Organic Carbon Content over Cultivated Areas: An Overview. Remote Sens. 14, 2917. https://doi.org/10.3390/rs14122917

Vaudour, E., Gomez, C., Lagacherie, P., Loiseau, T., Baghdadi, N., Urbina-Salazar, D., Loubet, B., Arrouays, D., 2021. Temporal mosaicking approaches of Sentinel-2 images for extending topsoil organic carbon content mapping in croplands. Int. J. Appl. Earth Obs. Geoinformation 96, 102277.

Vaudour, E., Gomez, C., Loiseau, T., Baghdadi, N., Loubet, B., Arrouays, D., Ali, L., Lagacherie, P., 2019. The impact of acquisition date on the prediction performance of topsoil organic carbon from Sentinel-2 for croplands. Remote Sens. 11, 2143.

# 5. Modelling land use in economic models at higher spatial scales

When coupling ABM with biophysical models, the two models need to be aligned in terms of scale and resolution. Many biophysical models are comprehensive about soil, climate and practices and have a high timely and spatial resolution (e.g. daily). This data is usually not available in ABM and therefore many assumptions need to be made.

The coupling of the BESTMAP ABMs with the biophysical models was complicated with the restriction in data access of the IACS/LPIS and the FADN datasets (Cord et al. 2023).

Sensitivity analysis of all models involved in linking can help to identify the most relevant sources of uncertainty and help to simplify the models by identifying the variables that can be fixed due to their low impact on the model output (Saltelli and Annoni 2010) .

Further research could involve the behavioural factors that investigate the decision about the specific piece of land that farmers choose to adopt agri-environmental schemes in order to provide spatially explicit information to the biophysical models.

### Linking to economic models

Most economic models focus on either the supply of food commodities and energy crops, or the demand for such products. This reflects the underlying dynamics encapsulated within the model, and ultimately whether they consider supply or demand to be the key determinant of economic activity:

- Neoclassical models are typically supply-driven, in that available inputs to production (including labour and capital) are combined via a production function to identify potential supply. In such a framework, land use represents a constraint on total maximum supply, in that it limits inputs to production.
- Keynesian models are demand-driven, and focus on drivers of demand (e.g. activity in the real economy determining demand for food and agricultural produce), with supply increasing to meet demand. Such models typically exclude land use considerations, alongside other limitations on the supply side of the economy.

As can be deduced from these descriptions, land use is more commonly incorporated into neoclassical supply-focussed models, where the supply constraints and trade-offs suggested by land use models can add substantial insight into the application of such frameworks. However, it is possible to build land use considerations into Keynesian models, as part of an effort to better represent supply-side constraints. For example, in the Horizon Europe project LANDMARC[1], Cambridge Econometrics (a member of the BESTMAP consortium) is developing the FTT:Agri sub-model.

FTT:Agri will be hard-linked to the macro-econometric E3ME model[2]. E3ME provides aggregate inputs into FTT:Agri, mainly the demand for food by around 20 different commodity types. FTT:Agri will simulate how the demand for food affects the substitution of crop cultivation in each of the 61 regions defined in E3ME. Based on that, food prices can be estimated, which feed back into the demand for food. The main strength of FTT:Agri (in conjunction with E3ME) is that it allows elucidation of the effects of policies on both the demand-side and the supply-side of food commodities and energy crops. Within the framework, it is possible to assess the highly debated competition for land between food crops and energy crops, and also between food crops and feed crops. FTT:Agri will therefore provide a novel and powerful approach to assess bio-energy supply in the context of the energy transition, food trade and the economic and environmental effects of future diets.

The model, as with the broader E3ME model, starts from the demand side of agriculture, estimating aggregate final use food demand for each region through a time-series econometric equation, disaggregated by product type, as well as agricultural product demand and domestic supply and trade. On the supply side, production yields for each crop type by region are estimated, alongside land capacity factors. The model then assesses intra-agricultural substitution of land use between product types, taking into account farmer preferences and exclusions to land-use substitution (e.g. unsuitable land), as well as intersectoral land-use substitution (i.e. the use of land for non-agricultural purposes such as housing, commercial buildings, industry and infrastructure). Agricultural land use substitution in this modelling framework is a function of preferences, which are assessed via levelized costs integrated into binary logit functions for each pair of

[1] See https://www.landmarc2020.eu/ for details
[2] See www.e3me.com

products, with the rate of substitution dependent upon the variation in profit margins and an uncertainty factor (representing imperfect knowledge). Farmers may decide to cultivate a different crop at some point in time. Such agricultural land use substitutions are dictated by what substitutions are permissible bio-physically speaking, and what the expected profit margins are for each crop. At the same time, the farmer does not possess perfect foresight of what will happen in the future and does not have access to all the information needed to make perfectly rational decisions about substituting one agricultural land use for another. This stems from the assumption of farmers being heterogeneous agents. An estimation of agricultural land use substitution based on economic considerations should capture such distributions around choices. The substitution metric is an adjusted Lotka-Volterra equation (commonly known as the predator-prey equation) (Volterra 1926, Lotka 1920).

Land-use change is also addressed in global CGE models, such as DART-BIO (Delzeit et al. 2018, Zabel et al. 2019) or MAGNET (Woltjer et al. 2014). The impact of land-use change to increase food production using either cropland expansion or intensification on endemness richness is addressed by e.g. Zabel et al. (2019) by linking a CGE model to a land-use model. The output of the latter is used to spatially assess impacts on biodiversity rich areas. Integrating non-market ecosystem services in models with a high spatial scale are rare since on national scales an economic value for non-market ecosystem services needs to be determined. A first approach is developed by La Notte (2020) by implementing natural capital accounts.

Note that this highlights just one potential approach to capturing land use decisions in macroeconomic models; there are a range of potential applications that can be developed, and each has relative strengths and weaknesses (for example, in the approach outlined above the strength is the treatment of uncertainty and heterogenous agents, which the weaknesses is the relatively low level of disaggregation of product types).

## Research needs for modelling agricultural policies in economic models

Market-based instruments like direct payments are well implemented in standard economic models. However, it is challenging to capture the impact of voluntary schemes using modelling tools such as partial equilibrium (PE) or general equilibrium (CGE) models. This challenge is relevant because voluntary schemes have become more important in recent amendments to the EU's agricultural policy.In the BESTMAP project, a new modelling framework has been developed to enhance the state-of-the-art modelling of the impact of agricultural policies on land use (see D6.5 for details). In the process of developing this modelling framework, several future research needs have been discovered.

### Background

To assess new agricultural policies, both local interactions of actors and market feedback at different spatial scales need to be considered to understand the policy impacts on nutrient losses, land-use change, food security, biodiversity and climate change (Kelly et al. 2013). While computable general equilibrium (CGE) or partial equilibrium (PE) models represent well market feedback, local interdependencies such as peer interactions due to social learning and land sales are better represented in agent-based models (ABMs) (Schmidt et al. submitted).

ABMs in the context of land use and agriculture are often used to link optimising farm models (e.g. Berger and Troost 2014; Möhring et al. 2016). Economic models such as PE/CGE models are hardly linked to ABMs, yet. However, many authors (Babatunde et al. 2017; Müller et al. 2020; Rounsevell et al. 2014; Niamir et al. 2020; Husby und Koks 2017; Huber et al. 2018; Millington et al. 2017), see the potential to represent bounded rationality and heterogeneity, but also to capture feedback with markets and to increase (spatial) resolution and scale. So far, there is only one experience with linking ABM with CGE/PE regarding land use (SWISSland). However, this link is no longer sustained due to model instabilities and the loss of expertise. In Millington et al. (2017), a concept for model linking is presented to represent "telecoupling", when land-use decisions on the local level have an impact on land use somewhere else. However, the authors used two "telecoupled" ABMs, instead (Dou et al. 2019). In Niamir et al. (2020), there is an example of linking an ABM with a CGE concerning consumer behaviour. However, this model is exploratory and cannot be applied for policy analysis as they base their model only on limited data from two NUTS 3 regions.

**Challenges encountered**

In the development of the modelling framework as well as during an expert workshop in Basel, six main challenges were identified which are described in detail in D6.5. ·

- Conceptual challenges
    - Alignment of different conceptual approaches
    - Different levels of aggregation
    - Model output interpretation and communication
- Technical challenges
    - High computational cost and challenges to linking variables
    - Model validation
    - High demand in the expertise of software development and funding

More specifically, we learned that many detailed data sources are under data protection and therefore not accessible for research. Different approaches exist to overcome these limitations such as synthetic populations or upscaling (see next section).

In addition, the workshop addressed in a sub-session the question: When does linking makes sense? This depends on the research question and on the topic at hand. It was concluded that is useful if different spatial scales need to be addressed and interact, e.g. local land-use decision and international trade flows or international prices. So, price changes from CGE/PE models could be transferred to ABM model to inform changing opportunity cost, while ABM could inform CGE/PE models about changes in behaviour due to social learning and changes in preferences, that would influence the elasticities of the CGE/PE models. Among the currently existing land-use models, PE models seem to be more suitable to link to ABM than CGE models. The direction of passing on information can be manifold but preconditions need to be fulfilled, e.g. matching sectoral and regional aggregations.

In the case of the BESTMAP models, the ABMs would have to consider different opportunity costs to make use of information from the CGE model. With the information from the current state of the models, information from the AMBs on adoption rates can be used in a PE model, but no link to a CGE model can be performed.

**Future research**
Several aspects need more research to enable the linking of typical economic models such as CGE/PE models with ABMs.

**Data availability**

With the method of synthetic populations, a new data set can be generated based on an aggregated variable and it is known or assumed distribution to imitate the heterogeneity of the data. The underlying data can be from several data sources (Pahmeyer et al. 2021). This approach is also known as disaggregation.

With the method of upscaling, disaggregated data from single farms is aggregated on the national or regional level. Based on certain characteristics such as age, farm type or income categories, the data is weighted to match the known national level (Zimmermann et al. 2015). However, there is only a limited amount of research on these two approaches and on their impact on model behaviour and the sensitivities of these methods on the model results. Therefore, to improve data availability, more research regarding synthetic population, aggregation and disaggregation is central.

Results of combined models become more reliable when there is more data from more case studies. Research is also needed on a minimum of agents to develop robust results.

**Conceptual alignment**
One key output of the expert workshop was that substantial additional and systematic work is needed to improve the understanding of when coupling makes sense and how to lay a suitable basis for it. This includes an improvement in communication between the different modelling communities since only a few research groups have experience in both model types. Further, in co-design with stakeholders, both modelling communities should ensure that the development of a linked model serves a useful research question.
Another conceptual alignment relates to the models' dynamics. The BESTMAP ABMs need more responsive opportunity cost to get dynamics in the model that respond to the price feedback from the CGE/PE models if the ABMs want to see the influence of (policy) shocks on the single farm.

—

**Using AES adoption rates in PE models**

One output of the BESTMAP ABMs is the adoption rate of AES for different farm types in the case study regions. In the PE model CAPRI, adoption rates of e.g. wheat coverage by farm is represented by a certain coefficient. A concrete application of improving a PE model is to replace these coefficients with the adoption rate of AES under different as simulated with ABMs.

## References

Babatunde, K. A.; Begum, R. A.; Said, F. F. (2017): Application of computable general equilibrium (CGE) to climate change mitigation policy: A systematic review. In: *Renewable and Sustainable Energy Reviews* 78, S. 61–71. DOI: 10.1016/j.rser.2017.04.064.

Berger, Thomas; Troost, Christian (2014): Agent-based modelling of climate adaptation and mitigation options in agriculture. In: *Journal of Agricultural Economics* 65 (2), S. 323–348.

Cord, Anna; Roilo, Stephanie; Beckmann, Michael; Paulus, Anne; Schneider, Katharina; Lugonja, Predrag et al. (2023): D3. 3 Ecosystem service, biodiversity and socio-economic models for each case study. In: *ARPHA Preprints* 4, e114653.

Dou, Yue; Millington, James D.A.; Bicudo Da Silva, Ramon Felipe; McCord, Paul; Viña, Andrés; Song, Qian et al. (2019): Land-use changes across distant places: design of a telecoupled agent-based model. In: *Journal of Land Use Science* 14 (3), S. 191–209. DOI: 10.1080/1747423X.2019.1687769.

Huber, Robert; Bakker, Martha; Balmann, Alfons; Berger, Thomas; Bithell, Mike; Brown, Calum et al. (2018): Representation of decision-making in European agricultural agent-based models. In: *Agricultural Systems* 167, S. 143–160. DOI: 10.1016/j.agsy.2018.09.007.

Husby, Trond G.; Koks, Elco E. (2017): Household migration in disaster impact analysis: incorporating behavioural responses to risk. In: *Natural Hazards* 87 (1), S. 287–305. DOI: 10.1007/s11069-017-2763-0.

Kelly, Rebecca A.; Jakeman, Anthony J.; Barreteau, Olivier; Borsuk, Mark E.; ElSawah, Sondoss; Hamilton, Serena H. et al. (2013): Selecting among five common modelling approaches for integrated environmental assessment and management. In: *Environmental Modelling & Software* 47, S. 159–181.

Millington, James D.A.; Xiong, Hang; Peterson, Steve; Woods, Jeremy (2017): Integrating modelling approaches for understanding telecoupling: Global food trade and local land use. In: *Land* 6 (3), S. 56.

Möhring, A.; Mack, G.; Zimmermann, Al; Ferjani, A.; Schmidt, A.; Mann, S. (2016): Agent-based modeling on a national scale–Experiences from SWISSland. In: *Agroscope Science* 30 (2016), S. 1–56.

Müller, Birgit; Hoffmann, Falk; Heckelei, Thomas; Müller, Christoph; Hertel, Thomas W.; Polhill, J. Gareth et al. (2020): Modelling food security: Bridging the gap between the micro and the macro scale. In: *Global Environmental Change* 63, S. 102085. DOI: 10.1016/j.gloenvcha.2020.102085.

Niamir, L.; Ivanova, O.; Filatova, T. (2020): Economy-wide impacts of behavioral climate change mitigation: Linking agent-based and computable general equilibrium models. In: *Environmental Modelling and Software* 134. DOI: 10.1016/j.envsoft.2020.104839.

Pahmeyer, Christoph; Schäfer, David; Kuhn, Till; Britz, Wolfgang (2021): Data on a synthetic farm population of the German federal state of North Rhine-Westphalia. In: *Data in Brief* 36, S. 107007. DOI: 10.1016/j.dib.2021.107007.

Rounsevell, M. D. A.; Arneth, A.; Alexander, P.; Brown, D. G.; Noblet-Ducoudré, N. de; Ellis, E. et al. (2014): Towards decision-based global land use models for improved understanding of the Earth system. In: *Earth Syst. Dynam.* 5 (1), S. 117–137. DOI: 10.5194/esd-5-117-2014.

Saltelli, A., & Annoni, P. (2010). How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, *25*(12), 1508-1517.

Schmidt, A.; Appel, F.;, Argueyrolles, R.; Baldi, L.: Filatova, T.; Finger, R.; Ge, J.; Grujić, N.; Heckelei T.; Huber, R.; Koç, A.A.; Li, C.; Mack, G.; Müller B.: Stepanyan D.; Will, M.;, Delzeit, R.; (submitted). Linking economic equilibrium models with agent-based Models. *Ecological Economics.*

Voinov, Alexey; Shugart, Herman H. (2013): 'Integronsters', integral and integrated modeling. In: *Environmental Modelling & Software* 39, S. 149–158.

Zimmermann, Albert; Möhring, Anke; Mack, Gabriele; Ferjani, Ali; Mann, Stefan (2015): Pathways to truth: comparing different upscaling options for an agent-based sector model. In: *Journal of Artificial Societies and Social Simulation* 18 (4), S. 11.

# 6. Data needs for biophysical modelling in agricultural landscapes

While the situation has slightly improved over recent years with the advent of novel datasets, such as those derived from remote sensing (e.g. COPERNICUS), the challenges associated with the lacking availability of data for biophysical modelling, especially in agricultural landscapes, persist. Despite advancements, there remains a significant gap in the accessibility and completeness of crucial data needed for accurate and comprehensive biophysical models. The complexities of agricultural systems demand detailed and up-to-date information on soil characteristics, weather patterns, land use changes, and crop dynamics. In many regions, the insufficient sharing of relevant data, the absence of standardised formats, and the reluctance of stakeholders to contribute information continue to hinder the development and refinement of effective biophysical models. Addressing these data limitations is paramount and should be prioritised.

Improved data availability extends benefits not only to the scientific community but also to farmers. For instance, having access to information detailing the yield impacts of transitioning to organic farming or applying Agri-Environmental Schemes (AES) empowers farmers to make informed decisions when contemplating such shifts. Additionally, comprehensive data on various crops used and their yields, especially when applied as Ecological Focus Areas (EFA) or within AES, proves invaluable. BESTMAP proposes several alleys for addressing these issues.

### Improving the usefulness of the Farm Sustainability Data Network (FSDN)

In BESTMAP, we used both agent-based modelling and biophysical modelling approaches, but the lack of individual data in FADN has made our approach challenging. To model the effects of farming practises on biodiversity and ecosystem services, it would be useful if FSDN data would include more detailed information on the land management. Therefore we suggest to add specific data requests to new versions of FSDN, including:

(1) land-use aspects, e.g. intensity, farming practices, pesticide application, fertiliser application
(2) agricultural yields per crop and area
(3) land tenure (owned/leased land)

### Improving access to spatial FSDN data

Presently, the management of data in HE projects faces considerable inefficiencies, resulting in both time and financial wastage. This is primarily attributed to intricate restrictions and requesting regulations, coupled with the intentional reduction of information, such as geographic locations, by the FADN/FSDN or IACS data holders. These limitations necessitate the development of elaborate workarounds, ultimately complicating the interpretation of model outcomes and, in some cases, rendering them devoid of meaningful insights. Addressing these challenges and streamlining data handling processes could significantly enhance the efficiency, clarity, and overall impact of HE projects.

There would always be data that FSDN cannot obtain from farmers, and models would be a powerful tool to augment FSDN, e.g. to get a better idea of soil carbon stored on agricultural land. However, any application of such models requires environmental data (e.g. climate, soil) from existing fine-resolution datasets. We believe researchers should be allowed access to anonymized micro-level FADN data that includes geographic locations of the fields each farm is managing, in order to conduct meaningful research using FSDN to make it less difficult "..for farmers, farm advisors and policy-makers to identify and implement sustainable approaches". Currently, due to the unnecessarily complicated and time-consuming process required to access the data many resources are wasted. Allowing access to micro-level anonymized data would be GDPR compliant as it could not be directly linked with individuals or companies.

### Improving compatibility of FADN/FSDN to other European data

FADN data collection represents a significant financial commitment and is a crucial source of data for research in the agricultural sector, but its use is limited due to missing spatial information and compatibility with other data. Specifically, FADN is not easily compatible with the IACS/LPIS data that are available for research in many EU member states, which include information on the location of individual fields (farmer's blocks), the farm that manages them and a range of ancillary data (crop type, AES adoption, etc.). Linking IACS/LPIS with micro-level FADN data is not possible because the latter represent a relatively small sample of farms from a large heterogeneous region (in some cases a whole country) without any information on their location. There are also significant differences in nomenclature and classification systems (e.g. of crops). An effort to harmonise those datasets, possibly by combining data collection efforts and saving costs, would make both data products better.

### Addressing issues related to the sampling / regions used in FSDN

There are a number of sampling issues in FADN which BESTMAP has been struggling with, and we would hope to see improved in FSDN:

First, some FADN regions are very large and internally heterogenous. For example, the whole of Czech Republic is a single FADN region which is quite limiting policy analysis within CZ. Having sau NUTS2 as the largest domain would help. Aligning FSDN to NUTS2 would also make it easier to link up FSDN statistics at regional level with many other EU surveys/ data products that operate at NUTS2 level.

Secondly, we would suggest making a better representation of organic farms. Being one of the Farm to Fork objectives, their lack of representation in FADN is hindering monitoring the implementation of this policy. As an example, this (DOI: 10.7251/agreng1603069w) says that there were 1.3% - 3% of FADN farms in Poland that were organic, and a Thünen Institute report from 2011 (https://orgprints.org/id/eprint/21012/1/D7_3_final.pdf) state: 'As can be seen from Table 4-1 only a few countries have a data set for organic farms which is big enough for analysis'.

Thirdly, FSDN data should make sure that smaller farms are also considered in a representative way. BESTMAP compared farm composition based on LPIS/IACS, FADN and FSS and found that a very significant proportion of farms are missing in FADN, although FADN does capture the majority of land area. Especially if the aim of FSDN is to also include “social sustainability” and rural development - we need information about the majority of land holdings. For instance for Germany there is a threshold of 25.000€ total standard output (see Annex I to Commission Implementing Regulation (EU) 2015/220): Taking data from 2016 this means that in Germany about 35 % of the farms recorded by the Federal Statistical Office are excluded from FADN (see also working paper on representativeness of FADN data).

Finally, BESTMAP is currently using ‘farms represented’ in FADN which is based on farm specialisation and economic size. This means it is impossible to upscale from FADN if you are asking questions about organic farms and intensive vs. non-intensive forms of farming. Having multiple weights or adding more dimensions to the weighting system would resolve this issue.

# 7. Piloting synthetic FADN generation

The growing global population (Ritchie et al., 2023) and the increasing urgency of ensuring food security under climate change (Schmidhuber & Tubiello, 2007), have sparked a dual transformative process in agriculture: the expansion of agricultural land-use (Ellis, 2011) and the intensification and industrialization of farming practices (Dias et al., 2016). To promote biodiversity conservation, it is crucial to embrace sustainable agricultural practices, with Agri-Environmental Schemes (AES) standing out as a prominent means through which governments can encourage such initiatives (Batáry et al., 2015a; Burton & Schwarz, 2013). Therefore, Understanding the factors influencing the adoption of AES is imperative for effective implementation. As part of the University of Leeds contribution to deliverable D5.5 we explored what affects the adoption of AES across holdings, countries, and years. We focus on 6 non-mutually exclusive hypothesis, 4 of which (H1-H4) relates to the holding’s level of organisation, while the remaining 2 (H5-H6) relate to the larger spatial and organisation-level scales at which the holdings operate:

- ***H1: Holding’s economic size and financial balance affects the adoption of AES.***
  Holdings with a larger economic size or in a favourable financial situation are more inclined to apply for subsidies (Coyne et al., 2021) since that may cover application costs (Falconer, 2000), and/or spare potential agricultural land for other usage.
- ***H2: Holdings characteristics such as area, yields, agricultural practices, natural resources, and owner/manager characteristics affect the adoption of AES.***
  Holdings with more intensified agricultural practices (e.g., higher yields) may be less likely to adopt AES, while holdings encompassing substantial sections protected under NATURA 2000 are more inclined to adopt AES (Paulus et al., 2022). The adoption of AES may also be influenced by farm ownership or management characteristics, e.g., younger owners (Morris & Potter, 1995).
- ***H3: Holding’s attitude towards environmental aspects affects the adoption of AES.***
  Reserving agricultural areas for biodiversity conservation, reducing farming intensity and investing effort into AES training hinges on a favourable stance towards environmental concerns (Batáry et al., 2015b; Bottazzi et al., 2018; McCracken et al., 2015).

- **H4: Holding's tendency to use governmental subsidies affects the adoption of AES.**
  Accessing information on rights, overcoming procedural barriers, and having trust in the government are essential prerequisites for applying to AES (Taylor & Van Grieken, 2015). Consequently, holdings that receive additional subsidies may exhibit a higher likelihood of applying for AES due to their enhanced capacity in navigating these requirements.
- ***H5: The environmental context of the holdings at the NUTS3 level affects the adoption of AES.***
  The environmental context of holdings may affect the feasibility and success of sustainable agricultural practices, influencing yields and subsequently defining the areas available for conservation (Beckmann et al., 2022).
- ***H6: The socio-economic context of the holdings at the NUTS3 level affects the adoption of AES.***
  Holdings in regions with low GDP per capita or a high percentage of poverty may be less inclined to adopt AES, as they may face the challenge of addressing more pressing and negative survival concerns.

The Farm Accountancy Data Network (FADN) is a comprehensive database that captures pivotal characteristics within the agricultural landscape and is thus highly suitable for exploring the above questions and hypotheses. However, the inclusion of personal and confidential information poses a challenge for widespread utilisation due to privacy concerns and legal considerations surrounding the protection of sensitive data. In navigating this delicate balance between widespread utilisation and confidentiality, the integration of synthetic data emerges as a promising technological solution, protecting sensitive data, improving the accuracy of machine learning models, and mitigating bias (Assefa et al., 2020; Figueira & Vaz, 2022; Nikolenko, 2019). We harvested the Synthetic Data Vault library (SDV)  comprehensive set of tools that covers the entire analytical pathway including data preparation, modelling, sampling, quality evaluations and visualisation (Patki et al., 2016) to generate synthetic data that mimics FADN data. Using the real FADN data and the synthetic data we focus on three main questions:

- Q1: Can we produce high quality synthetic data for the FADN database?
- Q2: What affects the adoption of AES across holdings, countries, and years?
- Q3: Will our conclusions from exploring Q2 with real data change if we explore it in a similar manner, but with synthetic data?

### Methods summary

The FADN database included 352 features across 28 countries and 4 years (2014-2017). The dependent variable was a binary version of AES subsidies ("SAEAWSUB_V" or "M_S_3300_V"), with all receiving holdings labelled as TRUE. FADN predictors included 58 features relevant to the four holding-level hypotheses (H1, H2, H3 and H4). In relation to H5 and H6, we collated 37 environmental features and 6 socio-economic features.

We generated synthetic data separately for each country-year combination with the SDV library version 1.8.0, in Python 3.11, (Python Software Foundation, 2024) for 52 features. After pre-treatment of data, we trained two synthesisers - the Fast ML Preset and CTGAN Synthesizer. The Fast ML Preset, a machine learning-based synthesiser, prioritises modeling time, ease-of-use, and preserving statistical properties (Fast ML Preset, n.d.). The CTGAN Synthesizer utilises GAN-based deep learning, employing a dual adversarial neural network system: generator neural network that create synthetic data competing against discriminator neural network that distinguishes between real and synthetic data (CTGAN Synthesizer, n.d.; Xu et al., 2019). The trained synthesisers were saved for future use (See SI files). Three sample sets were obtained from each synthesiser in each country and year: The mean number of holdings (2,962), the observed number of holdings and a constant of 5,000 samples. We assessed the quality of the generated synthetic data using various metrics from the SDMetric library (DataCebo, 2023), including: Key Uniqueness, Category Adherence, Boundary Adherence, Table Structure, column shape metrics (TV Complement and KS Complement), column pairs trends metrics (Contingency Similarity and Correlation Similarity), Category Coverage, Range Coverage and New Row Synthesis. For all metrics, a score of 1 means the synthetic data is similar in properties to the real data while a score of 0 means the synthetic data differs considerably from the real data.

Next, we applied Least Absolute Shrinkage and Selection Operator (LASSO) regression (Ranstam & Cook, 2018) of the dependent variable against the 58 FADN features (real and synthetic), 37 environmental features (real), and 6 socio-economic features (real). For synthetic data, we focused on samples generated by the CTGAN Synthesizer, using the observed sample size. LASSO analyses were conducted using seven predictor sets after stacking the data from all countries and years. The first set included all predictors, while the additional six sets represented hypotheses H1-H6. Each analysis utilised both real and synthetic data. The LASSO analysis was fitted with the glmnet R package version 4.1-8 (Friedman et al., 2010), in R version 4.3.2 (R Core Team, 2023), in two stages – tuning of the lambda hyperparameter and predicting for set-aside cases with 10 fold cross-validation. We assessed model quality on the set-aside fold with five indices: binomial deviance, misclassification error, mean squared error, mean absolute error, and AUC-ROC. The mean value over all folds was calculated for evaluation. We also compared the predicted coefficients between real and synthetic data analyses.

## Results and discussion:

### Q1: Can we produce high quality synthetic data for the FADN database?

All 672 generated synthetic datasets were identical to the real data in the most basic data structure attributes (e.g., column names, existence of unique keys, minimum and maximum range for numerical features, and levels for categorical and Boolean features). We generated a total of 2,446,930 rows of synthetic data and none of them was identified as identical to any of the original rows on which they were trained (see *New Row*

*Synthesis* metric in Figure 7.1). In other words, the data generated by both synthesisers was different enough from the real data to be considered unique.

However, the two synthesisers differed in their ability to conserve more complex structural attributes and correlations of the real data. In fact, the CTGAN Synthesizer outperformed the Fast ML Preset synthesiser in the column shapes and columns pair trends metrics (Figure 7.1). Higher columns shape scores mean that on average the frequencies of categorical levels in the synthetic data were more similar to the frequencies in the real data (*TV Complement, see FS.2 and FS.3*) and that the values of numerical features in the synthetic data were closer to the values in the real data (*KS Complement*). Higher columns pair trends scores means that on average the contingency tables between two categorical features in the synthetic data (e.g., all combinations of TF8 and ES6) are more similar to the contingency tables in the real data (*Contingency Similarity*) and that the Pearson correlation between two numerical features in the synthetic data is closer in value to the Pearson correlation between the same two numerical features in the real data (*Correlation Similarity*). For specific features, the CTGAN synthesiser was also better at covering the entire range of real data (*Range Coverage*, Figure 7.1). Both synthesisers were very good at ensuring all levels of categorical features are represented in the sampled data, even for features with high number of levels such NUTS3 (*Category Coverage*, Figure 7.1). Finally, The CTGAN synthesiser consistently outperformed the Fast ML Preset across different countries and years.

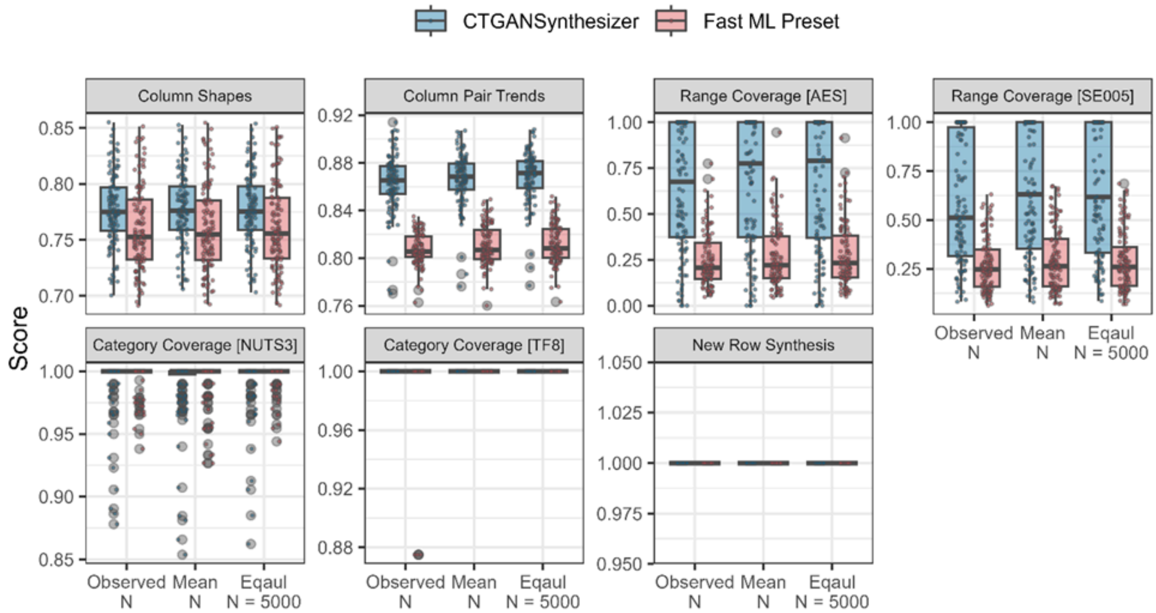


***Figure 7.1:*** *boxplots of the seven quality metrics quantified for each sample size and synthesiser. Each point is for a country/year dataset. AES is the value of Agri-environment subsidy [Euro] (feature SAEAWSUB_V). SE005 is the Economic size [Euro] feature. TF8 Is the Type of Farming with 8 categories.*

## Q2: What affects the adoption of AES across holdings, countries, and years?

LASSO models trained with all predictors outperformed all hypotheses-focused models (Figure 7.2). All performance measured identified the environmental context (H5, 37 features) and farm characteristics (H2, 22 features) as more informative, while holdings economic size and financial balance (H1, 24 features), holdings environmental attitude (H3, 8 features), holdings' tendency to use subsidies (H4, 9 features) and holdings socio-economic context (H6, 6 features) did not differ much in performance from one another and changed order in different metrics. However, the effect of the number of initial features on model performance is unclear and the hypotheses-focused models differed in that aspect.

The coefficient values themselves are not indicative of importance and since LASSO includes a built-in feature selection procedure (shrinkage of coefficients to 0), no additional variable importance index that compares the retained feature is available. In our application, feature selection resulted with 121 out of 125 features receiving non-zero coefficients. Therefore, it is impossible with the current analysis to identify the most important features – according to LASSO, they are all important. However, focusing on the coefficients of the type of farming (TF8) and economic size class (SIZ6) may provide some insights at least on these features since they are all binary and therefore their coefficients are relatively comparable to one another. For TF8 the largest coefficient was for 'other crops' holdings (0.766), followed by 'other grazing' (0.424). All other levels had a positive coefficient, except 'horticulture' holdings (-1.556). For economic size the base level was ES1 [2,000 – 8,000 Euro] and the coefficient increased with size from ES2 [8,000 – 15,000 euro] (0.778) to ES5 [100,000 – 500,000 euro] (1.337) and slightly decreased in the largest economic size ES6 [more than 500,000] (1.142).

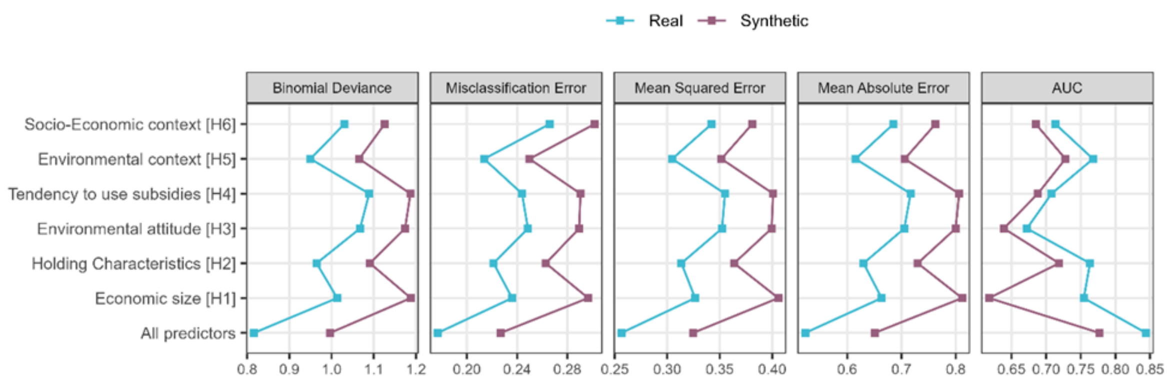


***Figure 7.2:*** *The mean over 10 folds of the performance metrics in LASSO analyses for the real and synthetic data, with all predictors, or with predictors relevant to each of the six hypotheses.*

## Q3: Will our conclusions from exploring Q2 with real data change if we explore it in a similar manner, but with synthetic data?

Despite the high metric scores (Figure 7.1), synthetic data utility is subjective and task-dependent, making it suitable for specific questions but inadequate for others (James et al., 2021). In the context of this study, we explored if our main conclusions about AES adoption would have looked different when explored with synthetic data. We found encouraging results (Table 7.1): the analysis based on synthetic data retained similar number of features, conserved the overall ranking of hypotheses-focused models (except for H1), produced coefficients that are strongly correlated to the coefficients of real data (Figure 7.3), conserved similar frequencies of features we consider important, and conserved high correlation between coefficients even within a single feature (Pearson r for TF8 and SIZ6). These, along with the decent scores from the SDV Quality Assessment and perfect scores for the *New Row Synthesis* metric (testing that all rows in the synthetic data are entirely unique), suggest that we may have found a good balance between information content and confidentiality.

However, not all details are kept. In the LASSO analyses, the models with real data outperformed synthetic data in all cases (Figure 7.2), suggesting that some of the complexity of interactions is inevitably lost.  In addition, when focusing on specific features, such as the economic size, the details matter. For the real data we observed an almost linear trend in coefficient values with economic size, and this information was lost in the synthetic data. For type of farming, the Pearson correlation was high as well (0.876, Table 7.1), and the coefficients of some of the levels were very close. However, for other levels the direction of coefficient changed, e.g., Wine holdings had a positive coefficient (0.255) in the real data, and a negative coefficient (-0.175) in the synthetic data. Another example is Hydrology of soil types (HOST): 11. HSG-A/D, which received the smallest coefficient with the real data (-11.463) and the second largest coefficient (4.815) with the synthetic data. Perhaps, such finer complexities may be retained with better synthetic data, produced by other synthesisers (e.g., Copula GAN synthesisers), or through different sampling strategies (e.g., conditional sampling). Furthermore, many modern statistical techniques thrive on large datasets which can be generated with ease using the synthesisers we have saved. Finally, the constraints inherent in the accounting nature of FADN may be better addressed if data synthesis occurred closer to the row data.

| Property | Real data | Synthetic data |
|---|---|---|
| | Feature selection | |
| Number of retained features | 121 out of 125 | 121 out of 125 |
| | Performance | |
| Best model (Figure 7.2): | All predictors | All predictors |
| Mean rank of performance | H5 [1} | H5 [1} |
| [Figure 7.2, | H2 [2} | H2 [2} |

| | | |
|---|---|---|
| 5 metrics, rank H1-H6, best model ranked 1] | H1 [3} | H6 [3.8} |
| | H6 [4.4} | H3 [4} |
| | H3 [5.2} | H4 [4.4} |
| | H4 [5.4} | H1 [5.8} |
| Visual inspection of Figure 7.2: | Lines are almost parallel | |
| Coefficients | | |
| Correlation between coefficients (Figure 7.3] | Strong positive correlation in all analyses | |
| Line of unity | Values close to line of unity | |
| Type of farming | | |
| Frequencies in data | Similar proportions in real and synthetic data | |
| Coefficients - Visual | Order not conserved, but some coefficients are close | |
| Coefficients - Pearson correlation | Pearson r = 0.876 | |
| Economic size class | | |
| Frequencies in data | Similar proportion in real and synthetic data | |
| Coefficient - Visual | Linear trend of coefficients from ES1 to ES5 in real but not in synthetic data | |
| Coefficients - Pearson correlation | Pearson r = 0.945 | |

***Table 7.1:*** *Contrasting the main result and conclusions from the LASSO analyses with real and synthetic data.*

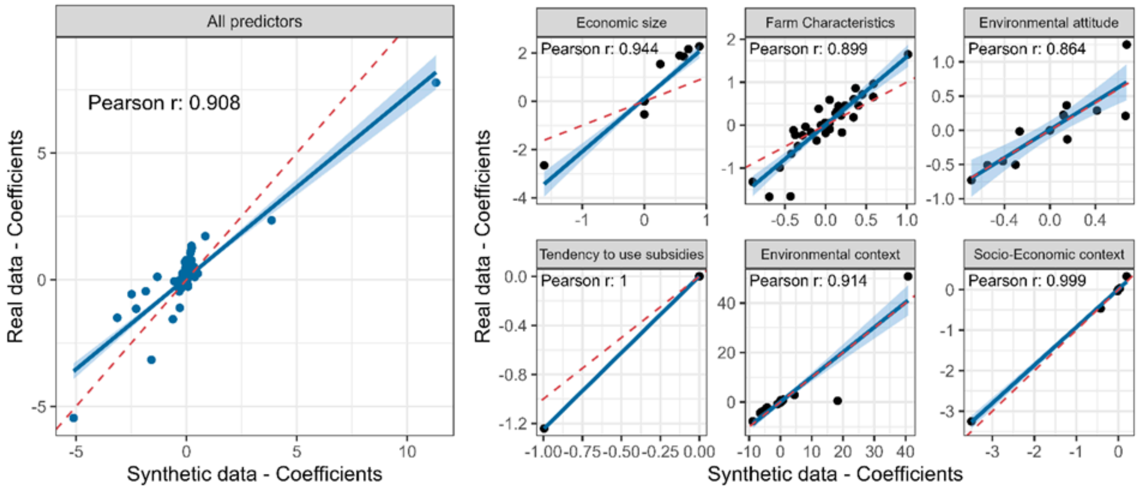


***Figure 7.3:*** *The LASSO features' coefficients in the analyses with the real data vs. the same features' coefficients in the analyses with the synthetic data for all predictors and for the predictors relevant to each of the six hypotheses. Each point represents a feature, and the blue line is a linear model fit of Real vs. Synthetic coefficients values. Pearson r correlation coefficients are given in each panel. Red dashed line is the line of unity.*

## Summary

Our generated synthesisers are available for use in Python and, to the best of our knowledge, are deemed safe for sharing as they do not retain the actual data in any form. Nevertheless, there exists a potential risk that the synthesised data may differ from the real data, albeit closely resembling it, thus raising concerns about potential disclosure risks (Abowd & Vilhuber, 2008) and broader-scale cyber threats. Prior to sharing, seeking professional advice to evaluate the safety of the synthetic data protocol and shared synthesisers is highly recommended. If found safe, our findings hint at a promising equilibrium, potentially striking a suitable balance between valuable information and robust confidentiality measures. As such, generation of synthetic data offers a unique opportunity for wider utilisation of the FADN database towards more sustainable agricultural practices and wider adoption of AES. In this study, we found that the adoption rate of AES may depend on multiple factors acting at different scales and capturing different aspects of the holdings. Therefore, unveiling the main factors affecting the adoption of AES may require analytical tools that can deal with such complexity. This study suggests that better understanding of such complexities may perhaps be achieved through synthetically generated big-data analysis.

## References

Abowd, J. M., & Vilhuber, L. (2008). How Protective Are Synthetic Data? In *Privacy in Statistical Databases* (pp. 239–246). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-87471-3_20

Assefa, S. A., Dervovic, D., Mahfouz, M., Tillman, R. E., Reddy, P., & Veloso, M. (2020). Generating synthetic data in finance. *Proceedings of the First ACM International Conference on AI in Finance*, 1–8. https://doi.org/10.1145/3383455.3422554

Batáry, P., Dicks, L. V., Kleijn, D., & Sutherland, W. J. (2015a). The role of agri‐environment schemes in conservation and environmental management. *Conservation Biology*, *29*(4), 1006–1016. https://doi.org/10.1111/cobi.12536

Batáry, P., Dicks, L. V., Kleijn, D., & Sutherland, W. J. (2015b). The role of agri‐environment schemes in conservation and environmental management. *Conservation Biology*, *29*(4), 1006–1016. https://doi.org/10.1111/cobi.12536

Beckmann, M., Didenko, G., Bullock, J. M., Cord, A. F., Paulus, A., Ziv, G., & Václavík, T. (2022). Archetypes of agri-environmental potential: a multi-scale typology for spatial stratification and upscaling in Europe. *Environmental Research Letters*, *17*(11), 115008. https://doi.org/10.1088/1748-9326/ac9cf5

Bottazzi, P., Wiik, E., Crespo, D., & Jones, J. P. G. (2018). Payment for Environmental "Self-Service": Exploring the Links Between Farmers' Motivation and Additionality in a Conservation Incentive Programme in the Bolivian Andes. *Ecological Economics*, *150*, 11–23. https://doi.org/10.1016/j.ecolecon.2018.03.032

Burton, R. J. F., & Schwarz, G. (2013). Result-oriented agri-environmental schemes in Europe and their potential for promoting behavioural change. *Land Use Policy*, *30*(1), 628–641. https://doi.org/10.1016/j.landusepol.2012.05.002

Coyne, L., Kendall, H., Hansda, R., Reed, M. S., & Williams, D. J. L. (2021). Identifying economic and societal drivers of engagement in agri-environmental schemes for English dairy producers. *Land Use Policy*, *101*, 105174. https://doi.org/10.1016/j.landusepol.2020.105174

CTGAN Synthesizer. (n.d.). *https://docs.sdv.dev/sdv/single-table-data/modeling/synthesizers/ctgansynthesizer*.

DataCebo, Inc. (2023). *Synthetic Data Metrics. Version 0.13.0 .* .

Dias, L. C. P., Pimenta, F. M., Santos, A. B., Costa, M. H., & Ladle, R. J. (2016). Patterns of land use, extensification, and intensification of Brazilian agriculture. *Global Change Biology*, *22*(8), 2887–2903. https://doi.org/10.1111/gcb.13314

Ellis, E. C. (2011). Anthropogenic transformation of the terrestrial biosphere. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1938), 1010–1035. https://doi.org/10.1098/rsta.2010.0331

Falconer, K. (2000). Farm-level constraints on agri-environmental scheme participation: a transactional perspective. *Journal of Rural Studies*, *16*(3), 379–394. https://doi.org/10.1016/S0743-0167(99)00066-2

Fast ML Preset. (n.d.). *https://docs.sdv.dev/sdv/single-table-data/modeling/synthesizers/fast-ml-preset*.

Figueira, A., & Vaz, B. (2022). Survey on Synthetic Data Generation, Evaluation Methods and GANs. *Mathematics*, *10*(15), 2733. https://doi.org/10.3390/math10152733

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1). https://doi.org/10.18637/jss.v033.i01

James, S., Harbron, C., Branson, J., & Sundler, M. (2021). Synthetic data use: exploring use cases to optimise data utility. *Discover Artificial Intelligence*, *1*(1), 15. https://doi.org/10.1007/s44163-021-00016-y

McCracken, M. E., Woodcock, B. A., Lobley, M., Pywell, R. F., Saratsi, E., Swetnam, R. D., Mortimer, S. R., Harris, S. J., Winter, M., Hinsley, S., & Bullock, J. M. (2015). Social and ecological drivers of success in agri‐environment schemes: the roles of farmers and environmental context. *Journal of Applied Ecology*, *52*(3), 696–705. https://doi.org/10.1111/1365-2664.12412

Morris, C., & Potter, C. (1995). Recruiting the new conservationists: Farmers' adoption of agri-environmental schemes in the U.K. *Journal of Rural Studies*, *11*(1), 51–63. https://doi.org/10.1016/0743-0167(94)00037-A

Nikolenko, S. I. (2019). Synthetic Data for Deep Learning. *CoRR*, *abs/1909.11512*. http://arxiv.org/abs/1909.11512

Patki, N., Wedge, R., & Veeramachaneni, K. (2016). The Synthetic Data Vault. *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 399–410. https://doi.org/10.1109/DSAA.2016.49

Paulus, A., Hagemann, N., Baaken, M. C., Roilo, S., Alarcón-Segura, V., Cord, A. F., & Beckmann, M. (2022). Landscape context and farm characteristics are key to farmers' adoption of agri-environmental schemes. *Land Use Policy*, *121*, 106320. https://doi.org/10.1016/j.landusepol.2022.106320

Python Software Foundation. (2024). *Python Language Reference, version 3.11. Available at http://www.python.org*.

—

R Core Team. (2023). *R: A Language and Environment for Statistical Computing. version 4.3.2.* R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>. .

Ranstam, J., & Cook, J. A. (2018). LASSO regression. *British Journal of Surgery*, *105*(10), 1348–1348. https://doi.org/10.1002/bjs.10895

Ritchie, H., Rodés-Guirao, L., Mathieu, E., Gerber, M., Ortiz-Ospina, E., Hasell, J., & Roser, M. (2023). Population Growth. *Our World in Data {https://Ourworldindata.Org/Population-Growth}*.

Schmidhuber, J., & Tubiello, F. N. (2007). Global food security under climate change. *Proceedings of the National Academy of Sciences*, *104*(50), 19703–19708. https://doi.org/10.1073/pnas.0701976104

Taylor, B. M., & Van Grieken, M. (2015). Local institutions and farmer participation in agri-environmental schemes. *Journal of Rural Studies*, *37*, 10–19. https://doi.org/10.1016/j.jrurstud.2014.11.011

Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). *Modeling Tabular data using Conditional GAN*.

# 8. Potential financing options for scaling up the approaches developed in BESTMAP

In the exploitation plan, a number of different potential funding streams are identified, specifically;

- Private companies
  - Food manufacturers
  - Food retailers
  - Think tanks
  - Agricultural producers
- Agricultural representative bodies
- NGOs and foundations
- Foundations
- Policymakers

While scaling up the approaches along the lines outlined in this roadmap would deliver substantially extended policy insights, it’s clear that substantial financial funding would be required to facilitate the development, and before any concrete conclusions (relevant either for the private or public sector) could be drawn. This is likely to rule out many of these identified funders as options. Private companies, even the largest agricultural firms, are unlikely to provide fundings for such a research-orientated purpose, while representative bodies and NGOs (and foundations) will similarly view the funding as not being relevant enough to their direct objectives (principally, influencing the direction of future agricultural policy) to be a relevant funding opportunity.

That leaves as potential options for providing finance a single option; public sector policymakers. Within this grouping, there are a number of potential financing streams that could be suitable;

- Research funding (e.g. Horizon Europe) - at both EU and Member State level, there are funds available dedicated specifically to the design and implementation of early- and mid-stage research projects. These could provide the required level of funding to extend the BESTMAP modelling approaches, although typically such research funds are i) very targeted in terms of themes, and ii) (depending upon the funding stream) highly competitive. As such, successfully accessing such financing would require a relevant call to be in place, that calls for the specific strengths of the BESTMAP designed approaches (and specifically, the extensions proposed in this roadmap). Specifically, the 'Food, Bioeconomy, Natural Resources, Agriculture & Environment' cluster under Horizon Europe Pillar 2, which focuses on Global Challenges and Industrial Competitiveness, is a potential research funding mechanism that could likely facilitate the development of an agricultural modelling framework like BESTMAP. Further research mechanisms are associated with a wide range of institutions and partnerships including European Research Council, Swiss National Science Foundation, German National Public Services, Strategy of German Ministry for Research and Education. Research for Sustainability, BMBF Research Initiative for the Conservation of Conservation Biodiversity, BiodivERsA.
- Policy analysis - while both the European Institutions (the European Commission and European Parliament) and Member State governments routinely commission quantitative policy analysis, to which the extended BESTMAP approaches could be applied, the scale of funding required to carry out the scaling-up work makes it relatively unlikely that Member States would put in the place the required funding. However, a long-term policy analysis contract (e.g. with DG Agri within the European Commission) could incorporate the upfront costs of the planned developments to the approaches, while also delivering scenario analysis to provide concrete insights into future policy development. In addition to the challenge of finding a call with the relevant amount of funding, a further challenge will be competing for such calls with existing established modelling frameworks. The best chance of success for such a call would come from a tender requiring the application of novel approaches (such as those developed in the BESTMAP project) to policy questions, which would recognise the specific strengths of the approaches that have been developed and provide scope for their expansion and application to addressing challenging policy questions.
- European Parliament - The European Parliament Research Service (EPRS) provides in-house research capacity to the European Parliament. As such, their expertise and knowledge covers a wide range of topics, including the application of modelling frameworks to policy questions. The Parliament has a specific Committee on Agriculture and Rural Development, and there is major scope for

new and innovative analytical work to feed into the deliberations of the AGRI Committee and therefore the policy-making process at the European Parliament. The combination of research capacity and policy making makes the EPRS and the European Parliament an ideal future user of the BESTMAP modelling framework, given the strong policy focus of the framework and the desire to inform policy making. Although it would be challenging for the EPRS to directly use the BESTMAP modelling framework without further guidance, potentially funding opportunities may arise from the European Parliament commissioning a guided use of BESTMAP outputs to inform the development of agricultural policies within the institution.

- Private sector finance - private companies, especially those in the agriculture and food sectors, could potentially provide funding. Despite the research-oriented nature of such projects may not align with their direct objectives, the BESTMAP consortium has hold numerous conversation with representatives of large agri-sector corporations in Europe (including Nestle and Danone) with the aim of providing them with the key information on the the BESTMAP modelling framework. Therefore, private sector organisations represent potential future funding opportunities with the aim of either developing the framework even further or commissioning a guided application of the BESTMAP outcomes to provide insights into the potential impacts of AEPs on ecological systems, and feed into the design of better AEPs in the future.

As can be seen from the above summary, due to the research-oriented nature of the BESTMAP outcomes there are only limited channels for accessing finance to scale up the existing BESTMAP approaches. The remaining dissemination work of the project consortium will be focussed on building relevant relationships with potential funders, by informing them of the unique properties of the BESTMAP modelling frameworks and the potential for expansion as summarised in this roadmap. Building these relationships, and ensuring a strong understanding of the BESTMAP approaches within key European and Member State institutions, is the most likely way of identifying and securing either research or policy analysis calls that can provide scope for further developing and scaling-up the approaches developed during the BESTMAP project.

# Appendix

Agenda for the “ABM/CGE modelling” workshop, May 12-13 2022, Basel:

## Program

Please fill in the Questionnaire for preparation by 26th of April

Location: Biozentrum, University of Basel, Spitalstrasse 41, 4056 Basel

https://unibas.zoom.us/j/65491710147?pwd=TEdTc0VUTWZSSWx5QmJSaktvTnNtdz09

Arrival from Germany via Badischer Bahnhof, Bus 30 ? Bahnhof SBB, Stopp: Kinderspital

Arrival from other destinations via Bahnhof SBB, Bus 30 ? Badischer Bahnhof, Stopp: Kinderspital

Day 1

| | |
|---|---|
| | --Coffee |
| 14:00-14:05 | Welcome by Guy Ziv (BESTMAP Coordinator) |
| 14:05-14:15 | Presentation of workshop goals (Ruth Delzeit) |
| 14:15-14:30 | Short introduction round |
| 14:30-14:45 | Input talks |
| | CGE and ABM linking for behavioral climate mitigation for consumers (Tatiana Filatova) |
| 14:45-15:30 | Presentation of different modelling approaches used in the agrimodels cluster |
| | AgriCore (Filippo Arfini) |
| | BESTMAP (Meike Will) |
| | MIND STEP (to be defined) |
| | --Coffee |
| 15:45-16:30 | Thematic parallel session: “How to combine ABM and PE/CGEs for specific research questions/policy domains” 1 /2 (specific domains will be defined based on the questionnaire) |
| 16:45-17:30 | Thematic parallel session: “How to combine ABM and PE/CGEs for specific research questions/policy domains” 3/4 (specific domains will be defined based on the questionnaire) |
| 17:35-17:50 | Summaries of the different sessions |
| 19:00 | Dinner |

Day 2

| | |
|---|---|
| 09:00-09:45 | Presentation of first ideas of paper outline and consequent identification of thematic parallel sessions |

| | |
|---|---|
| 09:45-10:30 | Thematic parallel session 5/6 |
| | --Coffee |
| 10:45-11:00 | Summaries of the different sessions |
| 11:15-12:30 | Road map for publication |
| | Lunch |

Accommodations:

The University of Basel has agreements for special offers for guests from the University of Basel for

- Hote Odeyla
- Hotel Spalentor
- Jugendherberge Basel
- Hotel Steinenschanze
- Hotel Krafft Basel

Basel is well connected to Weil am Rhein (D) and St. Louis (F), so you might also consider staying there.