

Modelling the Space-Time Dynamics of Soil Organic Carbon Stock: From Point to Landscape Scale

Yue Zhou

Thèse présentée en vue de l'obtention du grade de
Docteur en Sciences

**Université catholique de Louvain
Earth and Life Institute**



Jury members

Bas van Wesemael [Supervisor] (Professor, Earth and Life Institute, Université Catholique de Louvain)

Kristof Van Oost [Supervisor] (Professor, Earth and Life Institute, Université Catholique de Louvain)

Qiuzhen Yin [Chair] (Professor, Earth and Life Institute, Université Catholique de Louvain)

Veerle Vanacker (Professor, Earth and Life Institute, Université Catholique de Louvain)

Bertrand Guenet (Doctor of Science, CNRS at French National Centre for Scientific Research, France)

Fabio Castaldi (Doctor of Science, Institute of BioEconomy, National Research Council of Italy (CNR))

Acknowledgement

Several months ago, when I was imagining writing this acknowledgment, I thought I would feel very emotional. But now this moment has arrived, I feel surprisingly calm. These four years of my PhD seem to have passed in an instant. The scene of me leaving Beijing Airport four years ago is still vivid in my memory, and now it's already time to say goodbye again. During this short yet long time, I've created new memories and met many people I want to thank.

First, I want to thank my supervisor, Bas. I feel really lucky to be your student. You have deep wisdom and knowledge, and you always brought out inspiring ideas in our discussions. You are so supportive, revising our papers patiently many times over. Your humor filled our office with laughter. Most importantly, as Klara said, you are the kind of person "who is always there when you need him." Now it's time for you to rest, and I hope you enjoy a joyful retirement.

I also want to thank my co-supervisor, Kristof. In every discussion, you always brought thought-provoking ideas, prompting me to think more deeply and comprehensively about my work. I want to thank my jury members, Veerle, Fabio, Bertrand and Qiuzhen. Your constructive feedback helped fill the gaps and greatly improved my work.

Thanks also to Caroline and Klara for sharing code, and to Manon Ferdinand, Prof. Philippe Baret and Soil Capital for sharing the dataset on conservation agriculture and carbon farming. Additionally, thanks to Prof. Bernard Heinesch and Prof. Bernard Longdoz for sharing and explaining the dataset from the Lonzée station.

I would also like to thank the CSC scholarship (Grant No. 202006320061) for giving me the opportunity to study abroad and explore a broader world.

Of course, on this journey of exploration, I want to thank my friends for their companionship.

Thanks to the friends arrived in Belgium at the same time: Chenguang, Feilong, Jiadong, Shiqiang. We started our PhD under lockdown, figuring out how to get ID, bank accounts, and navigate various formalities. We took our first long trip abroad together, inexperienced and got scammed.

I would also like to thank my close peers, Lulu, Xiaojing, Yanfei, Yating, Xinrui. The time we spent together really brightened my life here. We cooked countless meals together, enjoyed dishes that I hadn't even tried in China, celebrated birthdays together every year, and created unforgettable photos and memories together across Europe.

Thanks to my short-term roommates, Jie and Lingmin. Living together felt like being with family and was truly heartwarming. And to my comrade Jie—I'm glad we both successfully completed our theses. To my colleagues at ELIC—Marmar, Huihong, Feba, Stephy, Pengzhi, Qianqian, Dries— thank you for the memorable times we spent together.

I am also grateful to my friends in China, Ruiying, Junsheng, Shuai, and others. Despite being miles away, always cared for me, believed in me, and encouraged me. Your companionship through video calls during my tough times gave me a lot of strength.

Thanks to my family for supporting me in every decision. I'm sorry I haven't been able to spend more time with them.

And to my grandfather—I guess you would be proud of me.

Louvain-la-Neuve, October 2024

Contents

Jury members	I
Acknowledgement	III
Contents	V
List of Abbreviations	IX
List of publications	XI
Chapter 1. General Introduction	1
1.1 Background	1
1.2 Modelling SOC spatial distribution	3
1.3 Modelling SOC temporal changes	5
1.4 Modelling SOC spatio-temporal changes	8
1.5 Key limitations	11
1.6 Problem statement	13
1.7 Thesis Objective	14
1.8 Workflow	14
1.9 Study Area	16
1.10 Thesis Outline	18
Chapter 2. Mapping SOC distribution	19
2.0 Outline	19
2.1 Introduction	20
2.2 Materials and Methods	23
2.2.1 Study area and soil samples.....	23
2.2.2 Environmental covariates:	25
2.2.3 Model Development	30
2.2.4 Empirical mode decomposition.....	35
2.3 Results	35
2.3.1 Descriptive statistics	35
2.3.2 Model performance	36
2.3.3 Importance of covariates	38

2.3.4	Mapping SOC content and its uncertainty.....	40
2.4	Discussion	42
2.4.1	Quality of the Prediction	42
2.4.2	SOC map products and its possible opportunities.....	43
2.4.3	Controlling factors of SOC patterns	50
2.4.4	Limitations and future perspectives	52
2.5	Conclusion	53
Chapter 3. Mapping conservation management practices		55
3.0	Outline	55
3.1	Introduction	56
3.2	Method and materials	60
3.2.1	Study area	60
3.2.2	Field data	61
3.2.3	Satellite data	64
3.2.4	Defining indicators for the three principles of CA.....	66
3.2.5	Model development and evaluation	74
3.3	Results	76
3.3.1	Species diversification.....	76
3.3.2	Accuracy for identifying cover crops and soil cover....	77
3.3.3	A model of non-inversion/inversion tillage practice prediction.....	78
3.3.4	Covariates selection and model accuracy for conservation agriculture field classification model.....	80
3.3.5	Classification map for conservation agriculture.....	81
3.4	Discussion	82
3.4.1	Track management practices with time-series imagery	82
3.4.2	Current status for classifying conservation agriculture	85
3.4.3	Potential applications and limitations of our work.....	86
3.5	Conclusion	87
Chapter 4 Simulating SOC Changes at Lonzée Station		89
4.0	Outline	89

4.1	Introduction	90
4.2	Materials and Methods	94
4.2.1	Site description	94
4.2.2	RothC Model	94
4.2.3	Boundary Conditions.....	95
4.2.4	Workflow.....	102
4.3	Results and Discussions	103
4.3.1	Changes in SOC stock.....	103
4.3.2	Uncertainties from different inputs into the model	107
4.3.3	Limitations of the work	112
4.3.4	Extensions and future work.....	113
4.4	Conclusions	114
Chapter 5 Evaluating SOC Changes at landscape scale		115
5.0	Outline	115
5.1	Introduction	116
5.2	Methods and Materials	119
5.2.1	Study Area	119
5.2.2	RothC model.....	121
5.2.3	Preparing the boundary conditions.....	122
5.2.4	Different spatial scales for the model runs	128
5.2.5	Model runs for different Management Scenarios	129
5.3	Results and Discussion	131
5.3.1	The effect of spatial scale on SOC dynamics.....	131
5.3.2	SOC Stock changes under different Scenarios.....	137
5.3.3	The effect of the choice for a baseline on calculating C sequestration.....	139
5.3.4	Limitations.....	142
5.4	Conclusions	144
Chapter 6. Summary and Conclusions		145
6.1.	Main Findings	145

6.2. Limitations and Perspectives	147
Appendix	150
Appendix 1 (Chapter 2).....	150
Appendix 2 (Chapter 3).....	151
Appendix 3 (Chapter 4).....	153
Appendix 4 (Chapter 5).....	159
Bibliography.....	161
List of Figures	189
List of Tables.....	193

List of Abbreviations

AGB	Above Ground Biomass
BAU	Business-as-Usual
CA	Conservation Agriculture
CAI	Cellulose Absorption Index
CC	Long-term Cover Crops
DEM	Digital Elevation Model
DSM	Digital Soil Mapping
EMD	Empirical Mode Decomposition
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization
FPAR	Fraction of Absorbed Photosynthetically Active Radiation
GBM	Gradient Boosting Machine
GEE	Google Earth Engine
IMF	Intrinsic Mode Functions
IPCC	Intergovernmental Panel on Climate Change
LAI	Leaf Area Index
LTE	Long-term Experiment
LPIS	Land Parcel Information System
ME	Mean Error
MRV	Monitoring, Reporting and Verification
MSE	Mean Squared Error
NBR2	Normalized Burn Ratio
NDTI	Normalized Difference Tillage Index
NDVI	Normalized Difference Vegetation Index
NPP	Net Primary Productivity
OA	Overall Accuracy
PA	Producer's Accuracy
RF	Random Forest
RFE	Recursive Feature Elimination
RMSE	Root Mean Squared Error
RT	Reduced Tillage
SAVI	Soil Adjusted Vegetation Index
SOC	Soil Organic Carbon

List of Abbreviations

SR	Surface Reflectance
STI	Simple Tillage Index
TA	Conventional Agriculture
TRI	Terrain Ruggedness Index
TWI	Topographic Wetness Index
UA	User's Accuracy
UUA	Utilized Agricultural Area

List of publications

Published

Zhou, Y., Chartin, C., Van Oost, K., van Wesemael, B., 2022. High-resolution soil organic carbon mapping at the field scale in southern Belgium (Wallonia). *Geoderma* 422, 115929

Qi, L., **Zhou, Y.**, Van Oost, K., et al. 2024. High-resolution soil erosion mapping in croplands via Sentinel-2 bare soil imaging and a two-step classification approach. *Geoderma* 425, 115930.

Wang, Z., **Zhou, Y.**, Zhao, R., et al. 2021. High-resolution prediction of the spatial distribution of PM_{2.5} concentrations in China using a long short-term memory model. *Journal of Cleaner Production* 284, 124731.

Submitted

Zhou, Y., Ferdinand, M., van Wesemael, J., Dvorakova, K., Baret, P., Van Oost, K., van Wesemael, B., 2024. A framework for mapping conservation agricultural fields using optical and radar time series imagery. Under review in *Remote Sensing of Environment*.

Chapter 1.

General Introduction

1.1 Background

Soil organic carbon (SOC) is a critical factor in terrestrial environments, acting as a primary energy source for microorganisms and influencing soil structure, fertility, ecosystem productivity, and the global carbon cycle (Lamichhane et al., 2019; Billings et al., 2021). Soil represents the largest carbon pool in terrestrial ecosystems, with approximately 1,500 Pg of carbon worldwide stored in the soil up to a depth of 1 meter. This is four times the carbon stored in the biotic pool and 3.2 times the amount in the global atmospheric carbon pool (Lal, 2004; Scharlemann et al., 2014). However, land cover changes and unsustainable land management during the Anthropocene, particularly intensive agricultural activities, have had detrimental effects on soil, leading to a continuous depletion of SOC stocks (Wei et al., 2014; Beillouin et al., 2023; Amundson et al., 2015). Sanderman et al. (2017) estimated that 31.2 Pg of carbon has been lost from surface soils (0–30 cm) over 12,000 years of agricultural practice. This has increased carbon dioxide (CO₂) levels in the atmosphere, thereby exacerbating the global greenhouse gas (GHG) effect (Bhattacharya et al., 2016; Chenu et al., 2019).

Till now, numerous studies have shown that increasing SOC stocks in agricultural land could significantly reduce the atmospheric CO₂ burden (Zomer et al., 2017; Smith et al., 2008; Pellerin et al., 2017). The “4 per 1000: Soils for Food Security and Climate” initiative, launched in 2015 (<http://4p1000.org>), aims increasing global SOC stocks by 0.4% annually. It is believed that a 0.4% increase in

SOC stocks in the top 1 meter of global agricultural soils has the potential to offset 20%–35% of anthropogenic greenhouse gas emissions (Minasny et al., 2017, Rumpel et al., 2020). Agricultural management practices can influence SOC stocks by either increasing carbon inputs to the soil or reducing SOC losses (Paustian et al., 2016; Chenu et al., 2019). Chenu et al. (2019) outlined several agricultural practices that can enhance SOC stocks (Fig. 1.1).

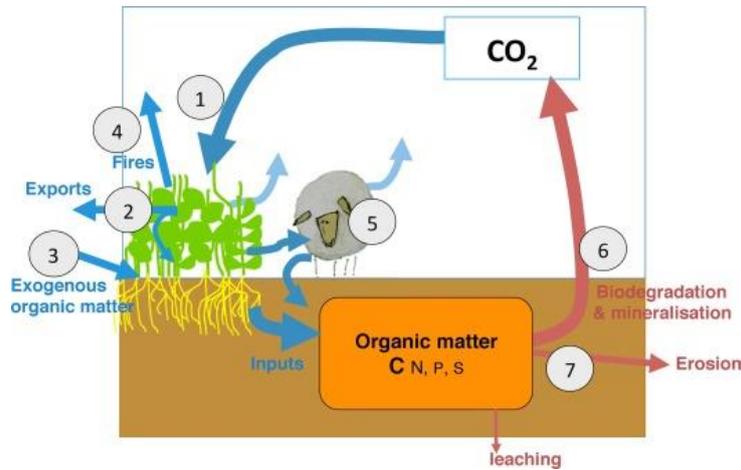


Fig. 1.1. Levers associated with agricultural practices that may influence SOC stocks: (1) increasing primary production (e.g. crop rotations, agroforestry, cover crops), (2) increasing biomass return to soil (crop residue return), (3) importing organic wastes to soil (manures, composts.), (4) avoiding fires, (5) grassland management (fertilization, grazing), (6) decreasing biodegradation and mineralisation rates (no tillage, water management), (7) decreasing erosion rates. (Chenu et al., 2019)

This potential of the soil to act as a carbon sink has fueled incentives for soil carbon sequestration, including policies such as emissions trading and soil carbon offset programs (Phelan et al., 2024; Barbato & Strong, 2023). For instance, the European Parliament recently approved the Carbon Removal and Carbon Farming (CRCF) Regulation (https://climate.ec.europa.eu/eu-action/carbon-removals-and-carbon-farming_en), certifying carbon removals, carbon farming, and carbon storage in products across Europe. Carbon farming rewards farmers for implementing climate-friendly practices, such as using cover crops, conservation tillage, catch crops, and improving

fertilizer use efficiency (Günther et al., 2024). It is important to note that while practices like importing organic waste to soil can enhance SOC stock and net primary productivity (NPP), they typically do not result in additional carbon transfer from the atmosphere to land, and thus are generally not considered true carbon sinks (Powlson et al., 2011).

However, the inconsistent sequestration rates across different management practices, soil types, and environments, combined with the high spatial variability of SOC stocks, contribute to significant uncertainty in sequestration estimates. (Luo et al., 2010; Gray et al., 2021; Stockmann et al., 2013; Padarian et al., 2022), which may be a barrier to the implementation of soil climate mitigation initiatives worldwide.

Evaluating the spatio-temporal dynamics of SOC is crucial not only for establishing carbon baselines for greenhouse gas emissions trading schemes and identifying and prioritizing potential locations for soil carbon sequestration projects, but also for uncovering local factors that influence soil carbon dynamics and for enhancing the monitoring and management of natural resources (Lamichhane et al., 2019).

1.2 Modelling SOC spatial distribution

The spatial variability of SOC is large, making it challenging to accurately characterize the spatial distribution of SOC based on a limited number of soil samples (Goodman & Owens, 2012).

The traditional approach is the **polygon-based soil map** (Lamichhane et al., 2019), which typically divides the study area into different soil map units (polygons of different soil types, land use types, vegetation types, geomorphological types, etc.), using the mean value of soil samples within each unit as the regional estimate. However, this method is heavily dependent on soil surveys, which are labor-intensive, time-consuming, costly, and rely significantly on expert knowledge (leading to substantial estimation differences based on varying partitioning methods). All the information is limited to the

sampling points, making it difficult to extend to large-scale applications (Chen et al., 2022; Pouladi et al., 2023).

Currently, **digital soil mapping (DSM)** is rapidly replacing outdated polygon-based soil maps and has become the most crucial method for obtaining spatial information on soil carbon (Pouladi et al., 2023; McBratney, 2016), especially at regional to continental scales (Padarian et al., 2017; Aitkenhead & Coull, 2016; Viscarra Rossel et al., 2014; Mulder et al., 2016). DSM builds the quantitative relationships between field or laboratory measurements with spatially explicit environmental data to predict the spatial distribution of soil properties (Fig.1.2).

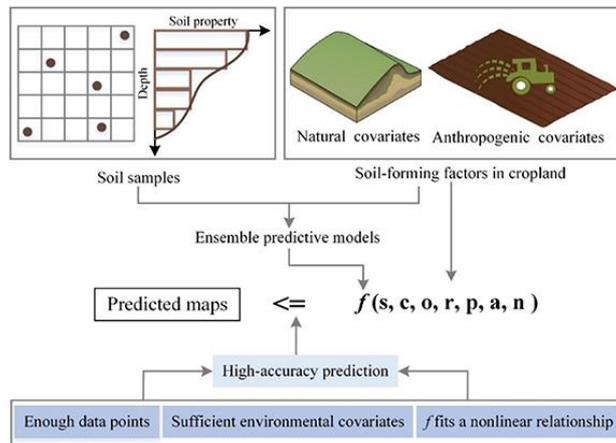


Fig. 1.2. Schematic diagram of DSM in cropland (Huang et al., 2022)

The foundation of DSM is rooted in the factorial methods introduced by Dokuchaev (1883) and later expanded by Jenny (1941). Jenny's conceptual model (CLORPT) posits that soil characteristics (S) in a landscape are a function of five environmental factors: climate (cl), organisms (o), relief (r), parent material (p), and time (t). While it has been useful in conventional soil mapping, it is neither quantitative nor spatially explicit (Kienast-Brown et al., 2022). To represent soil and the related environmental factors in a spatial context and express these relationships quantitatively, McBratney et al. (2003) further developed the Jenny's theory into the Scorpan model, which quantitatively describes the relationship between soil properties (S)

and environmental covariates (s, c, o, r, p, a, n) at a point in space and time as:

$$S = f(s, c, o, r, p, a, n)$$

The seven Scorman factors are soil (other soil information; s), climate (c), organisms (including anthropogenic factors; o), relief(r), parent material (p), age (a), spatial location (n).

The large-scale environmental covariates for DSM are typically obtained through remote sensing data from various platforms, such as unmanned aerial vehicles, airborne, and satellites (Pouladi et al., 2023). The SOC data collected from field or laboratory measurements can be linked to remote sensing data through different DSM techniques to predict and map SOC, including statistical, geostatistical, machine learning, and hybrid (combined) approaches (Chen et al., 2019). However, the correlation between SOC and environmental covariates may be influenced by image resolution, environmental conditions, and the scale of the study area (Lamichhane et al., 2019). These differences imply that the interpretation of SOC variability from remote sensing data must account for scale, as patterns observed at different scales can lead to varying conclusions (Wiesmeier et al., 2019).

1.3 Modelling SOC temporal changes

One method for assessing the temporal changes in SOC stock is based on the **soil monitoring network** (SMN, Morvan et al., 2008; van Wesemael et al., 2010; Saby et al., 2008a). This approach measures the temporal changes in SOC by revisiting the same locations, sampling, and measuring SOC contents. Many countries have implemented repeated soil sampling programs, such as France (Saby et al., 2008b) England (Bellamy et al., 2005), Germany (Peoplau et al., 2020), Denmark (Taghizadeh-Toosi et al., 2014), Belgium (Goidts & van Wesemael, 2007), China (Pan et al., 2010), and others. A challenge in resampling legacy data is that their locations were not originally selected with national survey purposes, raising doubts about their representativeness (Smith et al. 2019).

Additionally, the current SMN suffer from insufficient sample numbers. Morvan et al. (2008) analyzed the European SMN and determined that approximately 4,100 additional monitoring sites would be required to achieve a minimum sample density of 300 km² per site across Europe. For the existing SMNs, particularly in countries where the networks are relatively dense, it would take roughly a 10-year interval to detect significant changes in topsoil SOC contents (Saby et al., 2008a). National SMNs are typically more appropriate for assessing long-term trends in SOC accumulation or loss. However, when it comes to evaluating the impact of agricultural practices such as tillage or cover crops on SOC, the effect of these changes over 5 to 10 years is often too small relative to the total SOC pool (Beillouin et al., 2023; Fig. 1.3). Traditional sampling methods are therefore insufficient to capture these small changes over short periods.

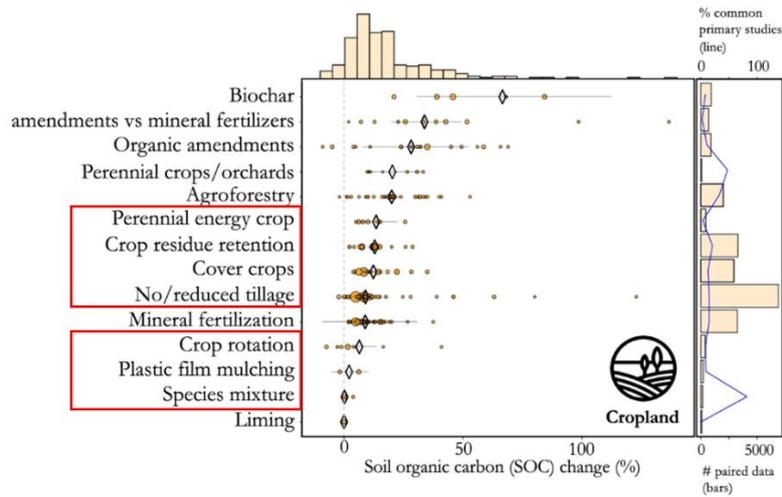


Fig. 1.3. Percentage change in soil organic carbon (SOC) due to land management practices in cropland (Beillouin et al., 2023)

To address this limitation, **process-based models** (also known as mechanistic models) can be employed to simulate temporal changes in SOC stocks. Compared to SMN, this approach significantly reduces costs and can also be used to extrapolate and predict SOC stock for unsampled years. These models can effectively simulate the impact of

management practices on SOC dynamics, providing valuable insights where SMNs fall short. Soil organic carbon models translate theoretical hypotheses into a simplified overview of the ecological system, described by schematic representations and mathematical equations (Le Noë, 2023). In simple terms, this involves quantifying SOC turnover using a series of mathematical formulas (Campbell & Paustian, 2015). Manzoni & Porporato (2009) identified approximately 250 models that are applicable to different temporal and spatial scales. These models focus on different level of interest (Fig. 1.4), from microbial community (microbiology/aggregate/rhizosphere), litter (plant residue decomposition), soil (soil organic matter dynamics), ecosystem (coupled soil–plant dynamics) to global (coupled model for global applications). Soil models, account for approximately 50% of the total, are predominantly used for relatively small spatial scales (up to the field level) and time scales ranging from daily to annual intervals.

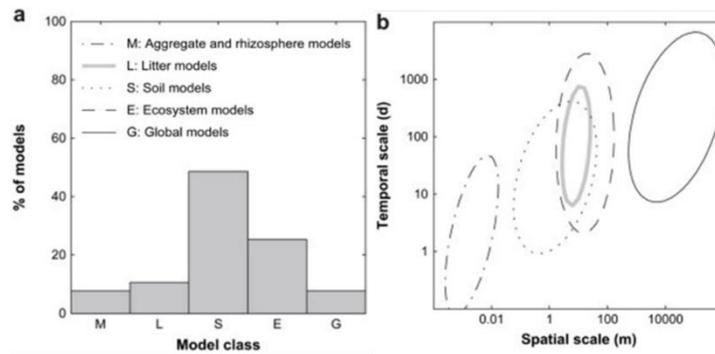


Fig. 1.4. (a) Percentage of models in different classes; (b) typical spatial and temporal scales of the model classes. The lines show the spatial and temporal ranges where 50% of the models in each class fall. (Manzoni & Porporato, 2009)

Most models are formulated with multiple pools using first-order decay kinetics (Fig. 1.5, Campbell & Paustian, 2015). The decay rate of this kind model is proportional to the SOC stock in the various pools considered, with rate modifiers implicitly representing key factors (such as soil temperature, soil moisture, and clay content) that influence microbial and physical processes. Since the 2010s, there has been a growing number of nonlinear kinetic models that represent the

decay rate as a function of SOC and/or microbial carbon pools to consider the feedback between microbial activity and SOC substrates (Le Noë, 2023). This kind of model thus enables the interpretation of soil microbial feedback responses to changing environmental conditions and SOC decomposition (Lawrence et al., 2009). Currently, an increasing number of Earth System Models (ESM) and Dynamic Global Vegetation Models (DGVM) incorporate these feedback effects to predict global carbon dynamics under climate change (Le Noë, 2023).

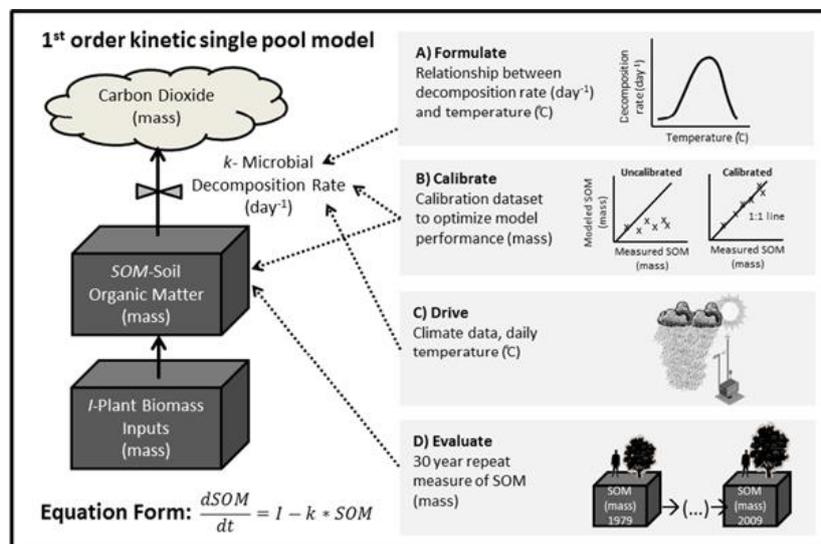


Fig. 1.5. An example of a single-pool soil organic matter dynamic model and the connection to hypothetical data used to formulate (A), calibrate (B), drive (C), and evaluate (D) model functions. (Campbell & Paustian, 2015)

1.4 Modelling SOC spatio-temporal changes

When we aim to simulate the spatio-temporal changes in SOC, there are two common approaches:

- 1. Projecting Spatial Simulations to Different Time Points**
(Spatio-temporal Digital Soil Mapping; DSM-ST)
- 2. Extending Temporal Simulations to Spatial Regions**
(Large-Scale Process-Based Models)

The process of DSM-ST involves first establishing static, empirical DSM models and then dynamically replacing environmental covariates to model SOC changes in both space and time (Heuvelink et al., 2020; Yigini & Panagos, 2016; Adhikari et al., 2019). However, a significant drawback of this DSM-ST approach is that SOC turnover is relatively slow, meaning the effects of driving factors on SOC changes do not immediately become apparent. This delay in the response of SOC to driving factors, when these factors are used in the DSM model for prediction, leads to time-lag effects when reconstructing long-term SOC time series (Xie et al., 2022). Additionally, DSM-ST relies on statistical relationships between SOC and covariates, without considering SOC turnover processes. This results in strong interannual fluctuations in the SOC time series due to the use of static covariates rather than dynamic processes (Xie et al., 2022).

In contrast, process-based models have the advantage of providing a clear mechanism to predict and understand more realistic SOC dynamics. First, DSM is used to estimate the initial soil state. Subsequently, large-scale climate, land use/land cover, and land management data are used to perform per pixel simulations across spatial domains, enabling the extension of SOC simulations from long-term experimental site scales to larger landscapes, regions, or even global scales (Morais et al., 2019; Jordon et al., 2022; FAO, 2020; Nadeu et al., 2015).

This workflow aligns with and supports the next steps in the monitoring, reporting and verification (MRV) framework (Smith et al., 2020; Fig. 1.6). It is important to note that the validation of SOC models (Fig. 1.6, item 3) can be conducted using data from short-term experiments (based on hourly/daily flux data; Fig. 1.6, item 2) or long-term experiments (based on long-term SOC measurements; Fig. 1.6, item 1). Given that this thesis focuses on the long-term effects of management practices on fields and that the changes brought by these management effects are small compared to large SOC stocks (Bai et al., 2019), we chose to use data from long-term sites for model evaluation.

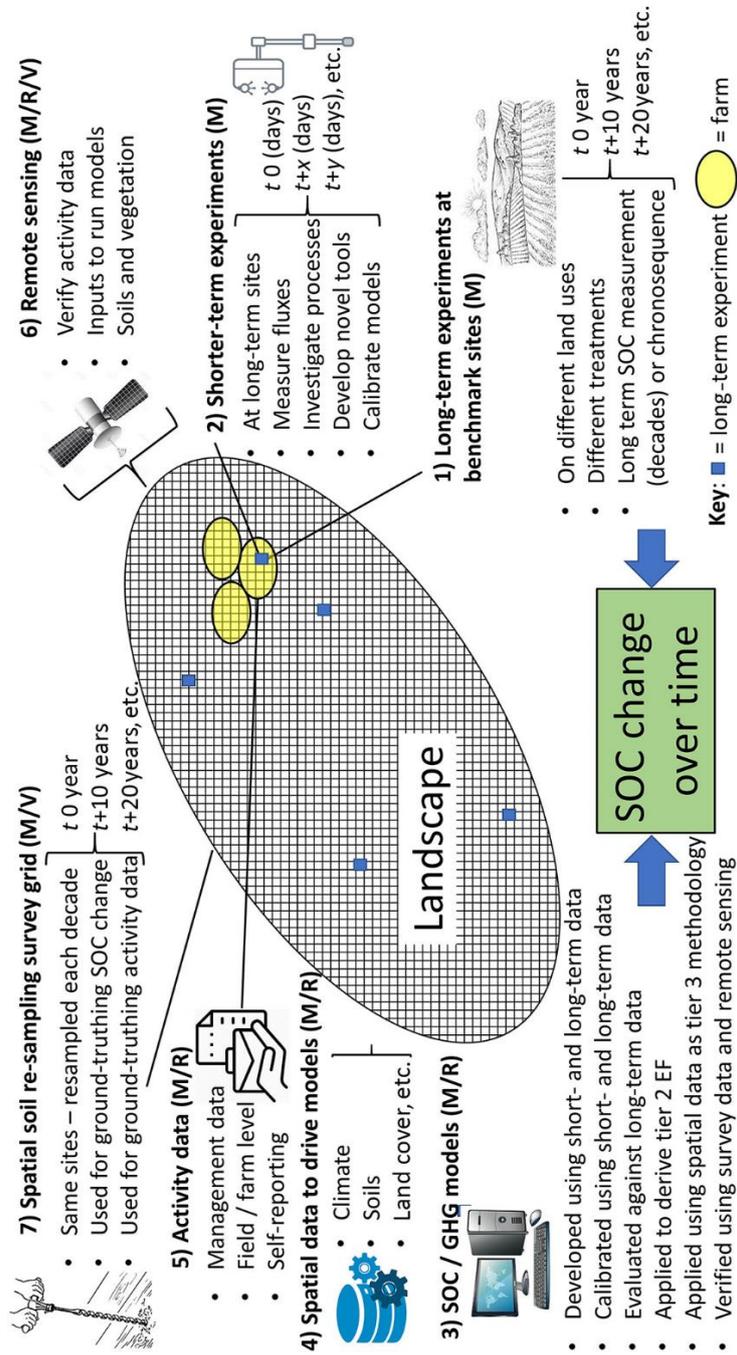


Fig. 1.6. Components of a soil measurement/monitoring (M), reporting (R) and verification (V) framework (Smith et al., 2020)

In addition, calibrating process-based models requires not only a time series of SOC data, but also, crucially, management data from long-term experiments (Fig. 1.6, item 5). This is especially important for croplands, where human activities significantly influence carbon inputs. For example, the widely used RothC model (Coleman & Jenkinson, 1996) requires data on farmyard manure (FYM) application and the carbon input from residues, which must be obtained from direct or indirect management records. However, a major challenge lies in the large-scale input data for models (Heuvelink et al., 2020), especially agricultural management data (e.g., organic fertilizer application, crop residue management, tillage practices, and cover crop measures), which are often coarse or scarce. However, due to the rapid advancements in remote sensing (Fig. 1.6, item 6) over recent decades, obtaining these management measures from remote sensing data is becoming increasingly feasible (Bégué et al., 2018; Kubitzka et al., 2020; Ahmed et al., 2024).

1.5 Key limitations

However, to date, several key limitations in spatio-temporal SOC prediction and carbon sequestration potential assessment still need to be addressed:

1. Covariates of DSM model:

Organism is one of the most important factors in the SCORPAN model, especially in the case of croplands, where human activities play a significant role, such as in the choice of crops, application of farmyard manure, and use of cover crops. However, current research often relies on average NDVI, NPP, or categorical variables (such as land use or vegetation type) to represent the organism factor (Padarian et al., 2017; Liang et al., 2019). These covariates are insufficient to capture the impact of human activities on SOC contents in agricultural fields. Additionally, few studies have demonstrated how covariates influence SOC distribution across different spatial scales.

2. Limitations in Large-Scale Management Data

Extending the process-based model to larger scales requires using remote sensing techniques to obtain regional boundary conditions. Current national or global-scale climate data are relatively accessible, such as the Climatic Research Unit dataset (University of East Anglia Climatic Research Unit, 2020), the ERA5 Monthly dataset (Bell et al., 2021), and the TerraClimate dataset (Abatzoglou et al., 2018). However, there is a significant lack of large-scale management data, particularly carbon input data, which is crucial for modeling. Regional simulations often have coarse carbon input estimates, typically derived from simple inverse mode runs (Jordon et al., 2022; Zhang et al., 2024) or the Miami net primary productivity (NPP) model based on temperature and precipitation (FAO, 2020; Poeplau & Dechow, 2023). In contrast, at long-term experiment (LTE) sites, more accurate carbon inputs are calculated using allometric functions combined with crop yield information (Keel et al., 2017). Therefore, if crop type and yield information at the field scale can be obtained regionally, similar calculations could be applied. The lack of organic fertilizer application, tillage, and other management data is also a major barrier to implementing large-scale process-based models.

3. Inadequate Agricultural Census Data:

A good foundation for this thesis is the open-access database provided by the Walloon Region of Belgium (<https://geoportail.wallonie.be/>), which includes information on time series of crops grown and field layouts (layouts i.e. the land parcel information system (LPIS)). However, the census database only contains data on main crops, not cover crops during the winter season. The main crop data combined with regional yield information, allows for more detailed carbon input estimates using allometric functions. However, cover crops can also contribute substantial carbon input (Poeplau & Don, 2015; Mazzoncini et al., 2011), ignoring this part could lead to significant errors. To address this, we need to use remote sensing data to predict management practices. Current methods for predicting cover crops typically focus on a fixed period during the winter months within a single year, usually from November to May. (Thieme et al, 2020; Barnes et al., 2021), which does not meet the

model's requirements for multi-year growth patterns, as well as duration of cover crop coverage.

4. Limitations in Defining Management Scenarios:

Due to the relatively rough estimates of carbon input in current work (Jordon et al., 2022; Zhang et al., 2024), which are not based on large-scale detailed management data, this uncertainty carries over to the step of simulating different management scenarios. Without knowing the source of carbon inputs (whether from manure, residues of main crop or cover crops), management scenarios can only be set as simple proportional adjustments to a baseline C input (FAO, 2020; Wiesmeier et al., 2016), which limits their practical value for developing or evaluating carbon farming strategies. For example, the Global Soil Organic Carbon Sequestration Potential Map (GSOCseq; FAO, 2020) defines three sustainable soil management practices corresponding to low, medium, and high carbon inputs, with increases of 5%, 10%, and 20%, respectively. However, it can not specify the management practices needed to achieve these increases. Only with data on actual management measures as a foundation, we can accurately improve management practices at the field level and more accurately assess the carbon sequestration potential under different scenarios.

1.6 Problem statement

SOC changes caused by management practices are small relative to the large and slow-changing SOC stocks (Beillouin et al., 2023), making direct sampling measurements challenging. An alternative is using process-based models that are already calibrated on long-term experimental sites. These models require intensive data inputs, which are not available for every field. Remote sensing can provide these input data, but the extent of uncertainty compared to measured data is unclear. Our hypothesis is that by producing reliable proxies and running the model with these proxies on long-term experimental sites, we can estimate the uncertainty of the models. This approach could then be applied to extensive areas while still identifying variability in management practices between fields.

1.7 Thesis Objective

The overall aim of this thesis is to simulate spatio-temporal changes in soil organic carbon stocks and sequestration under various management practices for cropland fields. The specific objectives are as follows:

1. Provide spatial data on soil conditions for process-based modeling.
2. Develop spatial datasets of management practices for process-based modeling.
3. Evaluate the effectiveness of readily available spatial data by comparing it with well-documented long-term experimental data to assess simulation uncertainties.
4. Apply the developed methods to a case study for monitoring, reporting, and verifying carbon farming practices.

1.8 Workflow

The detailed workflow for achieving the specific objectives is as follows: We use $t=0$ to represent the base period, and $t=1$ to denote the end period of the SOC modeling.

1. Develop a spatial model for SOC distribution:

Based on soil sampling data, obtain environmental covariates, especially those related to management, to build a DSM model and predict SOC spatial distribution in the base period (Fig. 1.7).

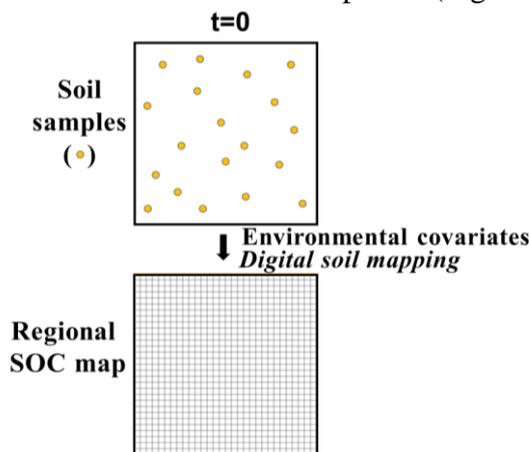


Fig. 1.7. Predict SOC spatial distribution at $t=0$

2. Simulate SOC changes over time using the RothC model:

Validate the RothC model using long-term experimental site data from Belgian cropland, based on monthly recorded climate and management data (Fig.1.8).

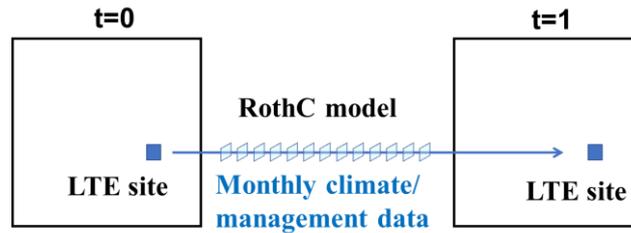


Fig. 1.8. Simulate SOC change over time using RothC model with long-term experiment (LTE) data.

3. Predict large-scale time-series climate and management data:

Providing an alternative from in-situ to remote sensing for all climate and management data, with a particular focus on preparing regional information on fertilizer application, yield, cover crops, and tillage practices (Fig.1.9).

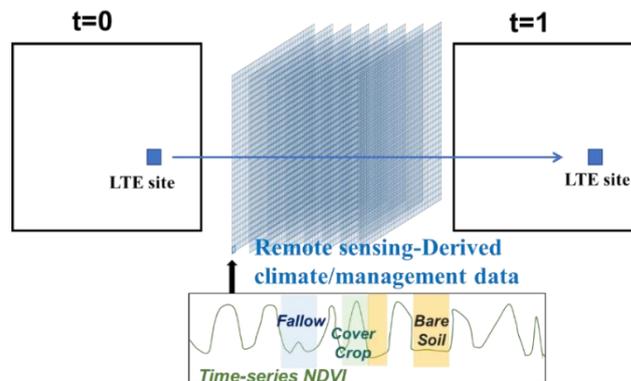


Fig. 1.9. Prepare large-scale climate and management data using remote sensing, e.g. time-series NDVI could be used to analyze presence of cover crop.

4. Run RothC model with time-series spatial data:

Based on DSM SOC map at $t=0$, using large scale boundary conditions, run the model from $t=0$ to $t=1$ (Fig.1.10).

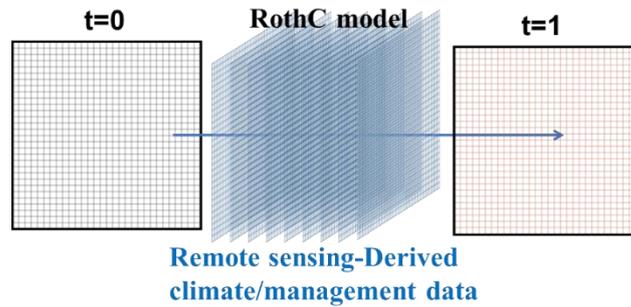


Fig. 1.10. Predict SOC spatial distribution at $t=1$.

5. Create management scenarios to explore their impacts on SOC Sequestration:

Investigate how changes in management practices affect SOC sequestration (Fig.1.11).

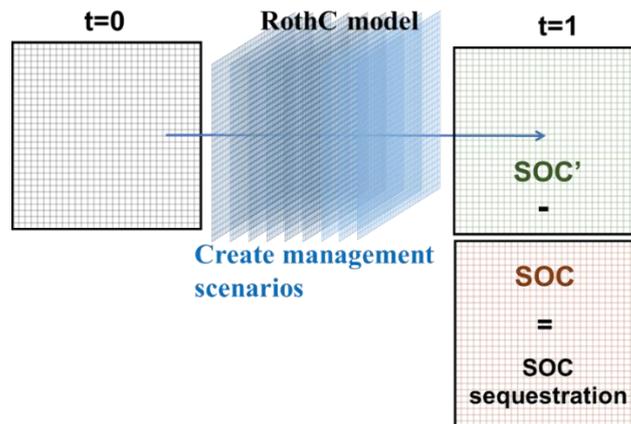


Fig. 1.11. Calculate SOC Sequestration between different scenarios.

1.9 Study Area

Our study area is located in the southern part of Belgium, specifically in the Walloon region. Geographically, Wallonia lies at approximately $50^{\circ}50'N$ latitude and $4^{\circ}00'E$ longitude in Europe (Fig. 1.12), and is divided into nine agro-geographical regions (Fig. 1.12 background). These regions are classified based on biophysical conditions of the natural environment and the similarity of agricultural

systems, as described by the Walloon geoportal (<https://geoportail.wallonie.be/catalogue/3ec1510b-7e87-4f92-ab8a-22675249d84b.html>).

Soil data (Chapter 2) and management data (Chapter 3) were collected across this diverse region, allowing for a comprehensive assessment of agricultural practices. The soil predicted map in Chapter 2 covers most of the Wallonia region, while the prediction of management practices in Chapter 3 is focused on the Hesbaye region, where is an agriculturally advanced area with fertile silt loam soils supports major crops such as cereals, sugar beet, potatoes, and flax. In Chapter 4, the model was run at the Lonzeé Terrestrial Observatory, located at 50°33'05.7"N, 4°44'46.1"E in Belgium, recognized for having one of the longest and most complete data series on cropland in Europe. Finally, in Chapter 5, all available data were integrated to complete a landscape-scale case study.

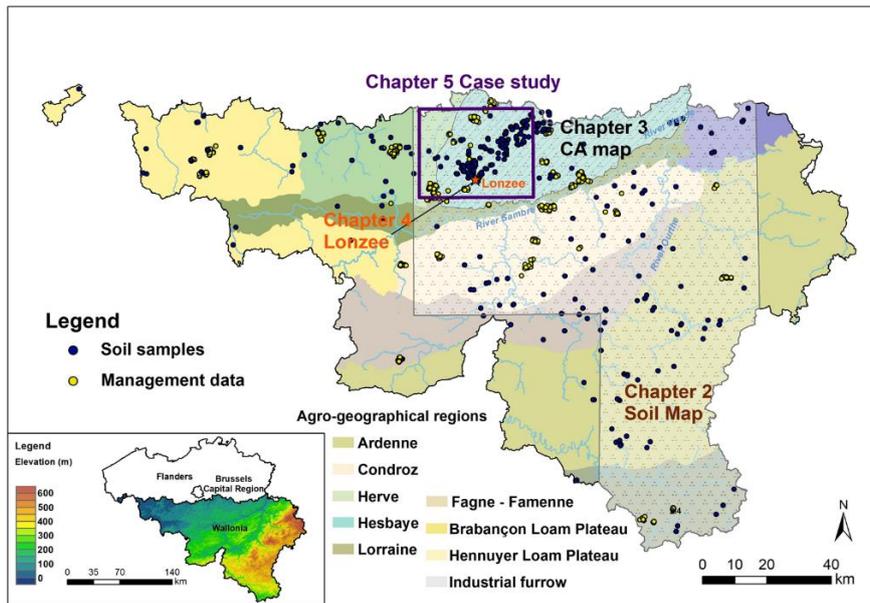


Fig. 1.12. Study area for each chapter of the thesis.

1.10 Thesis Outline

General introduction of this thesis	Chapter 1
We use the DSM approach to predict the regional SOC distribution for the base period. Prepare environmental covariates and primarily focus on human influence factors. We also explore and analyze how different environmental covariates impact the spatial distribution of SOC at various spatial scales.	Chapter 2
Due to the complexity of predicting management practices at large scale, we first analyze this question before modeling temporal SOC changes. In this chapter, we use long-term series remote sensing data to establish several indicators and identify management practices such as cover crops, crop rotation and tillage practice.	Chapter 3
Using dataset from the long-term experimental site in Belgium (Lonzée Terrestrial Observatory), we simulate SOC changes at the Lonzée site from 2007 to 2017, comparing the accuracy and uncertainty of using in-situ versus remote sensing boundary conditions on model simulations.	Chapter 4
We extend the model to the landscape scale, simulating changes in regional SOC stocks and calculating the carbon sequestration potential under different management scenarios.	Chapter 5
Conclusions, limitations and perspectives	Chapter 6

Chapter 2.

Mapping SOC distribution using Digital Soil Mapping

The content of this chapter was published in Zhou, Y., Chartin, C., Van Oost, K., van Wesemael, B., 2022. High-resolution soil organic carbon mapping at the field scale in southern Belgium (Wallonia). Geoderma 422, 115929.

2.0 Outline

Accurate soil organic carbon content estimation is critical as a proxy for carbon sequestration, and as one of the indicators for soil health. We collected 497 soil samples during 2015 and 2019, as well as five environmental covariates (organic carbon input from the crops, normalized difference vegetation index (NDVI), elevation, clay content and precipitation) at a resolution of 30 m. We then aggregated these to represent agricultural fields and compiled a soil organic carbon (SOC) content map for the agricultural soils of Wallonia using Gradient Boosting Machine. We calculated OC input from both main crops and cover crops for each individual field. As the cover crops do not occur in the agricultural census, we identified cover crops based on long time-series of NDVI values obtained from the Google Earth Engine platform. The quality of the SOC predictions was assessed by validation data and we obtained an R² of 0.77. The Empirical Mode Decomposition (EMD) indicated that OC input and NDVI were the dominant factors at field scale, whereas the remaining covariates

determined the distribution of SOC at the scale of the entire Walloon region. The SOC map showed an overall northwest to southeast trend i.e. an increase in SOC contents up to the Ourthe river followed by a decrease further to the South. The map shows both regional trends in SOC and effects of land use and/or management between individual fields. The field-scale map can be used as a benchmark and reference to farmers and agencies in maintaining SOC contents at an appropriate level and optimizing decisions for sustainable land use.

2.1 Introduction

Soils constitute the largest terrestrial carbon pool and contain about two to three times more carbon than the atmosphere (Batjes, 1996; Houghton, 2007). At the same time, soil organic matter is a main factor in aggregate stability, water retention and mineralization of nutrients (Robinson et al., 2012; FAO, 2015). Given the significant amount of carbon stored in the soil, a small change in the soil organic carbon (SOC) pool could exert significant impacts on the atmospheric CO₂ concentration (Smith, 2008). The Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report stated that soil carbon sequestration is the main mechanism responsible for 89 % of the agricultural greenhouse gas mitigation potential (Smith et al., 2007). However, injudicious land use and management of agricultural land could result in decrease of SOC contents and a vicious circle of soil degradation and soil threat (Lal, 2015; Wiesmeier et al., 2018). In this context, an explicit and reliable quantification and spatialization of SOC content is of great importance for soil quality assessment, especially as a baseline for assessing the effect of regenerative agriculture.

Soil property maps are increasingly becoming available because of digital soil mapping (DSM) initiatives throughout the world (Minasny and McBratney, 2016; Chen et al., 2022). DSM is an approach relying

on multivariate analysis between SOC and influential environmental covariates to construct soil maps using legacy soil data with multivariate regression, geostatistical and data mining methods (McBratney et al., 2003). The Scorpan model (McBratney et al., 2003), as an extension of Jenny's equation (Jenny, 1941), is the core of DSM and can fit quantitative relationships between SOC contents and seven Scorpan factors (i.e. soil (other soil information), climate, organisms (including anthropogenic factors), relief, parent material, age and position) (Chen et al., 2018). The *GlobalSoilMap* initiative (Arrouays et al., 2014) is at the forefront of such efforts, with the ambitious goal of mapping soil attributes at a 90 m resolution in six standard depth layers. A large number of countries have already contributed to *GlobalSoilMap* (McBratney et al., 2003; Adhikari et al., 2013; Padarian et al., 2017; Aitkenhead & Coull, 2016; Viscarra Rossel et al., 2014; Mulder et al., 2016; Liu et al., 2022).

Such maps with broad geographic extent highlight patterns of soil properties such as SOC and are useful for assisting decision making at national and regional scales. Nevertheless, they do not provide enough detail for farmers in order to allow them to restore their most degraded soils or to be used in the framework of carbon accounting (Malone et al., 2017, 2018). The legacy data does not usually reflect the current SOC contents, and the 90 m grid cells are not fine enough to distinguish spatial heterogeneity within and between individual fields and finally, the effects of historical management at the field scale do not appear. In order to represent organism factors when modelling SOC, both continuous variables derived from temporal series of satellite images such as mean NDVI and net primary productivity (NPP), and categorical variables such as land use or vegetation types are used (Padarian et al., 2017; Liang et al., 2019). However, these covariates are not sufficient to deconstruct

the influence of anthropogenic activities (e.g. crop rotation, cover crops, tillage) on SOC.

For a more precise reconstruction of crop and land management factors, precision farming uses remote sensing usually based on dedicated proximal and remote sensing platforms (e.g. Internet of Things, Unmanned Aerial Vehicles, Unmanned Ground Vehicles (Weiss et al., 2020). Although hyperspectral remote sensing has proven its success in field-scale SOC mapping (Gomez et al., 2008; Castaldi et al., 2018), the technology is still too expensive and often restricted to small pilot zones. Recently, the increased free accessibility to remote sensing images (Sentinel 1 and 2, Landsat, Gaofen, etc.) at fine resolution allows deriving multiple environmental covariates. Meanwhile, easy access to data and emerging cloud computing platforms (Google Earth Engine, GEE) helps us to synergistically use these satellite data faster and simpler (Weiss et al., 2020). The rapid acquisition of multiple time series of remote sensing images offers the possibility of describing the crop phenological evolution (Amin et al., 2021), cropping intensity (Liu et al., 2020) and the crop types (Piedelobo et al., 2019) of a single field. In parallel to this increase in available data, development of machine learning methods and deep learning algorithms also open the way for mapping soil properties at a finer resolution (Gholizadeh et al., 2018; Vaudour et al., 2019; Bousbih et al., 2019; Lin et al., 2020; Kumar et al., 2017; Castaldi et al., 2019). However, most of these studies set spectral bands as covariates, and even though a good accuracy can be obtained, it is difficult to understand and explain the inherent mechanisms and relationships between bands and SOC content.

The overall aim of this study is to create SOC maps under the DSM framework with high-resolution remote sensing data and derive new covariates (OC input from crop residues), that can distinguish the SOC content both at a field scale relevant for management decisions

and at the regional scale relevant for monitoring SOC contents as a proxy for soil health and agricultural greenhouse gas emissions. Simultaneously, we evaluate the covariates together with the scale at which they influence the SOC distribution. Meanwhile, with the aid of the interactive geospatial platform GEE and R, the SOC modelling procedure is optimized to the extent that it can be expanded to wider geographical areas in the future. Hence, trustworthy SOC maps will provide farmers with more detailed information about their fields and can be used as a baseline map for their decisions on sustainable soil management.

2.2 Materials and Methods

2.2.1 Study area and soil samples

We focus on the southern part of Belgium, known as the Walloon Region (Wallonia). Wallonia is a landlocked region with a total area of 16,901 km². The elevation gradually increases from north-west to south-east together with the precipitation from 800 to 1400 mm, while the temperature (from 10 to 8 °C) shows an opposite trend (Fig. 2.1). A change from deep sandy loam and silty soils to shallow silt loam and stony soils occurs in the same direction, along with a shift from intense arable agriculture to more extensive cattle breeding (Chartin et al., 2017; Goidts and van Wesemael, 2007). The Meuse and its tributary the Sambre are rivers that divide Middle Belgium into a northern and southern part. The climate north of these rivers is marine, whilst the south of this axis is more continental.

A total of 497 soil samples were collected between 2015 and 2019 (Fig. 2.1 (A)), each composite soil sample (including 5 subsamples) was obtained within a circle of 4 m radius circle (the microsite) centred on the soil profile, recording the geographical coordinates and sampling date, as well as the land use type (cropland and grassland)

and some generalities about the management practices for each field. In cropland, the soil was sampled in the homogeneous till depth while in grassland was sampled around 15cm. The samples were first air-dried, crushed, and passed through a 2-mm sieve, the total carbon content was determined by dry combustion using a VarioMax CN Analyzer (Elementar Analysensysteme GmbH, Hanau, Germany). The carbonate content of the samples that showed clear reactions with 10% HCl were determined using a modified pressure-calimeter method (Sherrod et al., 2002), the SOC content was then calculated by subtracting the inorganic carbon content from the total carbon.

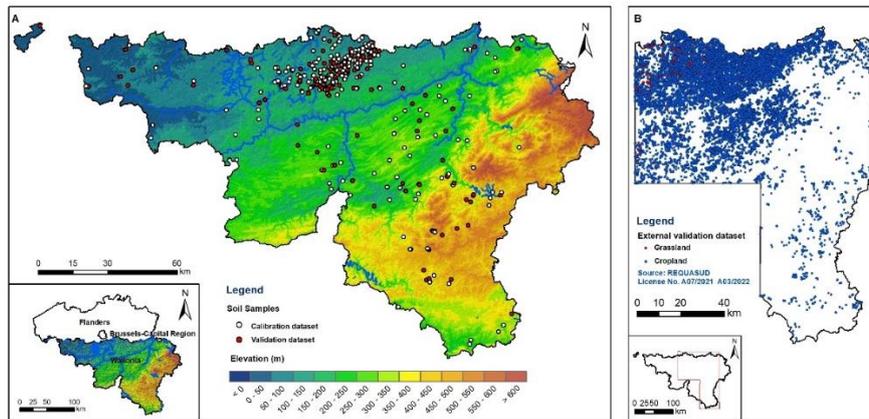


Fig. 2.1. Digital elevation model of Wallonia and locations of 497 soil samples (A) (70% calibration samples in white and 30% validation samples in red) and 21,551 external validation data (B)

An external validation dataset of 21,551 SOC content records (2015-2019) was provided by REQUASUD (License No. A07/2021 & A03/2022, Fig. 2.1 (B)), a network of Walloon laboratories. The soil samples were collected to a depth of 25 cm in cropland and of 15 cm in grassland in a standard “W” shape per field, and were then thoroughly mixed to form a composite sample. Samples were air-dried, crushed and sieved at 2 mm. SOC content was analyzed using both the Springer Klee method relying on oxidation of organic matter by potassium dichromate in a hot acid environment according to ISO

14235 and the dry combustion method according to ISO 10694, these two methods were considered as equivalent after inter-lab comparisons on specific samples. We did not use a soil depth function to harmonize soil data because the vertical distribution of SOC under deep inversion ploughing is relatively homogeneous (Priori et al., 2024).

2.2.2 Environmental covariates:

This study proposed a covariate to quantitatively characterize the OC input from crop residues of the main crop and cover crop. The procedure can be described as:

1. Obtain OC input from main crops (section 2.2.1)
2. Obtain OC input from cover crops (section 2.2.2)
3. Calculate the total OC input (section 2.2.3)

Please refer to section 2.2.4 for the acquisition of other environmental covariates

2.2.2.1 OC input from main crops

We collected the data on the main crop sown (2015-2019) from the website of the Walloon Region (<https://geoportail.wallonie.be/>, Request No. 1612978328110 & No. 1612734897423). The main crop allotment information (include grassland) is updated annually and recorded as polygons based on declarations from farmers, on-site verification, or remote sensing checks.

The organic C input to the soil for each crop (expressed in $\text{kg}\cdot\text{C}\cdot\text{m}^{-2}$) as well as the humification constant were extracted from the literature (Table A.1; Soil Service of Belgium & Ghent University, 2006). We considered the effective carbon input from the crops, contributing to the buildup of SOC, to be the product of the OC input and the humification coefficient.

2.2.2.2 OC input from cover crops

Unfortunately, the website of the Walloon Region only specifies the main crop and does not provide information on the cover crop. We obtained time-series of NDVI values for a five-year interval for each field, removed the period during which the main crop grows based on the crop calendar, then identified whether cover crops are present or not based on the NDVI values from the remaining period. As it was not possible to determine the type of cover crop, we only obtained the number of winters during which cover crops were planted over the past five years. We then calculated the mean OC input for a cover crop based on the data for the cover crops typical for Belgium (Table A.1). A detailed description of the procedure is as follows:

1. Obtain time-series of NDVI values from Google Earth Engine (Fig. 2.2): For the given study area, a defined time period (2015.01.01-2019.12.31) and the cloud cover (<85%), three datasets were initially filtered out. The collection included Landsat 7 Surface Reflectance (SR) dataset (391 scenes), Landsat 8 SR dataset (433 scenes) and Sentinel 2 SR dataset (1707 scenes).

After a more stringent cloud removal based on the image bands (detailed process in script path: <https://code.earthengine.google.com/8fea476933a1238e3c20afad6c4251d5>), the following processes such as reprojection, resampling and co-registration were performed (Nguyen et al.,2019; Claverie et al., 2018). As there is a slight difference between the red and NIR bands between Landsat 7, Landsat 8 and Sentinel 2, a band correction was applied. Finally, the time-series of NDVI values for the five-year period were exported in CSV format (Ujaval, 2020).

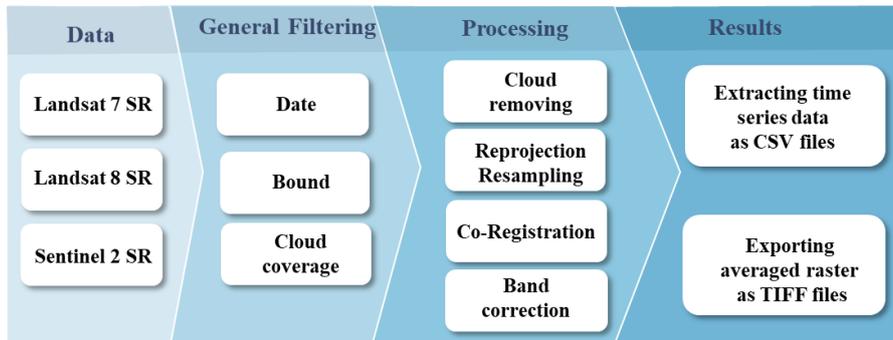


Fig. 2.2. Workflow of extracting time-series NDVI values in the Google Earth Engine.

2. Remove NDVI value for the period during which the main crop grows (based on crop types and crop calendar): Multiple data sources were collected (German Weather Service, www.dwd.de) to define a crop calendar for each crop type. Only the start time of sowing and the date of harvest were required, then we removed the NDVI data for the period between these two dates.

3. Judge the type of crop (summer/ winter): reclassify all crops to summer and winter crops, cover crops can only be grown during the winter preceding a summer crop.

4. Calculate mean NDVI value for the period without a main crop, if mean NDVI bigger than 0.4, count 1, else count 0, finally the sum represents the number of cover crop seasons during the five-year period.

Fig.2.3 shows a typical example for the identification of cover crops, where the green line reflects the NDVI values for one field during the 2015-2019 period. After obtaining the information of crop types and the corresponding crop calendar, five grey background blocks were added to the figure. For the remaining periods, if the mean values during the cover crop growth window were above 0.4, then it is identified as cover crop, and a green background block was added.

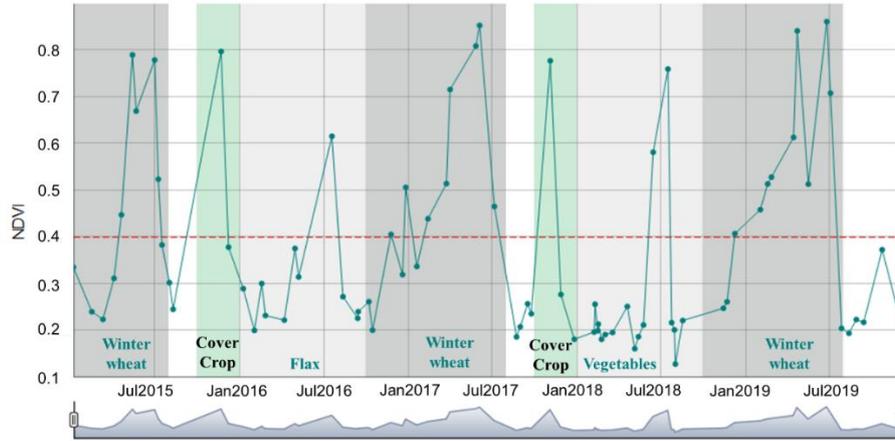


Fig. 2.3. Typical example of the NDVI time series in a single field used to illustrate the identification of the frequency of cover crops ($n=2$).

2.2.2.3 Total OC input calculation

We calculated the total OC input from growing crops using the biomass input of each main crop, each cover crop and their respective humification constants, then averaged the summation using the years (Eq. 2.1)

$$OC_{input} = \frac{(\sum_{i=1}^n OC_i \times HC_i) + OC_{cc} \times HC_{cc} \times T_{cc}}{n}, \quad (2.1)$$

where \bar{OC}_i is the supply of OC from crop residues, and HC_i is humification coefficient for year i ($i=1, 2, \dots, n$). T_{cc} is the number of cover crop during n years. OC_{cc} and HC_{cc} are the averaged OC and HC of cover crop.

2.2.2.4 Other covariates

In addition to the OC input, we collected a total of 15 covariates for agricultural land in the Walloon region that responded to the Scorpan model (Table 2.1).

Table 2.1. Covariates used for modeling SOC contents. (R: Resolution)

Factors	Covariates	R	Source
Terrain	Elevation (DEM)	30 m	Google Earth Engine
	Slope		SAGA GIS
	Aspect		
	Curvature		
	TRI (Terrain ruggedness index)		
	TWI (Topographic wetness index)		
Organism	NDVI (Normalized difference vegetation index)	30 m	Google Earth Engine
	LAI (Leaf Area Index)		
	FPAR (Fraction of absorbed photosynthetically active radiation)		section 2.2.1~2.2.3
	OC input (Organic carbon input)		
	Land use types	Discrete	https://geoportail.wallonie.be
Climate	Temperature Precipitation	40 m	Chartin et al., 2017
Soil	Clay	40 m	Chartin et al., 2017
	Silt		
	Sand		

Terrain information at a resolution of 30 m was collected from the shuttle radar topographic mission (SRTM, NASA /USGS/JPL-Caltech). We obtained the Digital Elevation Model (DEM) data from GEE and calculated other DEM derivatives (including slope, aspect, curvature, terrain ruggedness index (TRI) and topographic wetness index (TWI)) in SAGA (System for Automated Geoscientific Analyses) GIS.

The NDVI during 2015 to 2019 were obtained through the GEE platform. The process is described in detail in the section 2.2.2. The

only difference is in the step of exporting the results (Fig. 2.2), where the five-year NDVI values were averaged and produced as raster data with a resolution of 30 m. Moreover, we computed the Fraction of absorbed photosynthetically active radiation (FPAR) using NDVI data and Leaf Area Index (LAI) using soil adjusted vegetation index (SAVI) data. The specific formula and steps can be found in Allen et al. (2007) and Sellers et al. (1996).

Based on ground meteorological stations data from Belgium, Netherlands, Germany, France and the Grand Duchy of Luxembourg, precipitation and temperature was predicted at a resolution of 40 m using thin-plate splines regression, with elevation introduced as an auxiliary variable in the modelling. The smoothing parameters were selected automatically by generalized cross-validation (Fig. 2.4C; Chartin et al., 2017).

From 1949 to 1965, the National Soil Survey (NSS) of Belgium was carried out, with over 13,000 soil profiles described and analyzed. A digitized legacy database known as "Aardewerk" was compiled (Van Orshoven et al., 1988), which contains soil texture data. Using 6129 soil profiles from this database, Chartin et al., 2017) developed soil texture maps (40 m, including sand: 50 μm –2 mm, silt: 2–50 μm , clay: < 2 μm) by regression kriging (Fig. 2.4D).

Data for land use types (cropland and grassland) were also obtained from the website of the Walloon Region (<https://geoportail.wallonie.be/>, Request No. 1612978328110 & No. 1612734897423). All raster layers of environmental covariates were masked by agricultural region and resampled to 30m grid, then the averaged values were calculated for each field.

2.2.3 Model Development

2.2.3.1 Principle of Gradient Boosting Machine

Gradient Boosting machine (GBM) is a highly flexible and powerful machine learning technique which was originally derived by Friedman (2001). It can cater to many particular data-driven task (Natekin & Knoll, 2013), and has shown superior predictive capacity and considerable success in both data-mining challenges and soil property modelling (Pittman and Brown, 2011; Johnson & Zhang, 2014; Mishra et al., 2020). GBM produces a regression or a classification model in the form of an ensemble of a series of weak and inaccurate learners. In an incremental, additive and sequential manner, GBM consecutively fits new models to minimize the loss function and thus provide a more accurate estimation of the response variable.

The GBM framework provides flexibility, practitioners can design loss functions and choose several hyperparameter tuning options, which is one of the greatest advantages of GBM (Natekin & Knoll, 2013). GBM has been shown to outperform random forest in predict time and accuracy when parameters are carefully tuned (Chong, 2022; Sharma, 2015). But at the same time, the high flexibility leads to many parameters interacting with each other and heavily influencing the prediction. This increases the importance and difficulty of parameter adjustment.

The implementation of GBM in R can be realized with *gbm* R package. We constructed a grid of hyperparameter combinations (including shrinkage, depth of interaction, trees terminal nodes, fraction of bagging, trees), and performed a grid search to iterate every combination of hyperparameter values (Table A.2) rather than manually tweaking hyperparameters one at a time. Based on the minimum RMSE of the out-of-bag (OOB) data, we could retrieve our optimal combination of hyperparameters and train the subsequent models (UC Business Analytics R Programming Guide, 2018).

2.2.3.2 Model Calibration and Validation

Conditioned Latin Hypercube Sampling (CLHS, Roudier et al., 2012) is an efficient stratified random sampling method which can cover the entire range of the ancillary variables (Minasny & McBratney 2006). Based on the coordinates (Latitude and longitude) and SOC contents of the samples, CLHS was employed to split the data into a calibration and validation dataset with a similar spatial and SOC distribution. Thus, the soil samples (n=497) were partitioned into calibration dataset (70%, n=348) and validation data (30%, n=149).

Using the initial optimal combination of parameters derived from section 2.2.3.1, the model was first developed using all covariates ('full model') with calibration data, the model performance was evaluated by 10-fold cross-validation by means of three commonly suggested indices: (the determination coefficient (R^2), the root mean squared error (RMSE), the mean error (ME) and Ratio of Performance to Deviation (RPD) see. Eq. 2.2-2.5).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (2.2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - f_i)^2}{n}}, \quad (2.3)$$

$$ME = \frac{\sum_{i=1}^n (y_i - f_i)}{n}, \quad (2.4)$$

$$RPD = \sqrt{\frac{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2}} \quad (2.5)$$

where \bar{y} is the mean of the measured data, and y_i and f_i are measured and predicted values for sample i ($i=1, 2, \dots, n$) respectively.

Then, we removed covariates one by one from the 'full model' and measured new simplified models' accuracy, and deleted these

covariates if there was no change or a gain in model accuracy. At the end, precipitation, OCinput, NDVI, DEM and clay were chosen as the final model input covariates (Fig. 2.4).

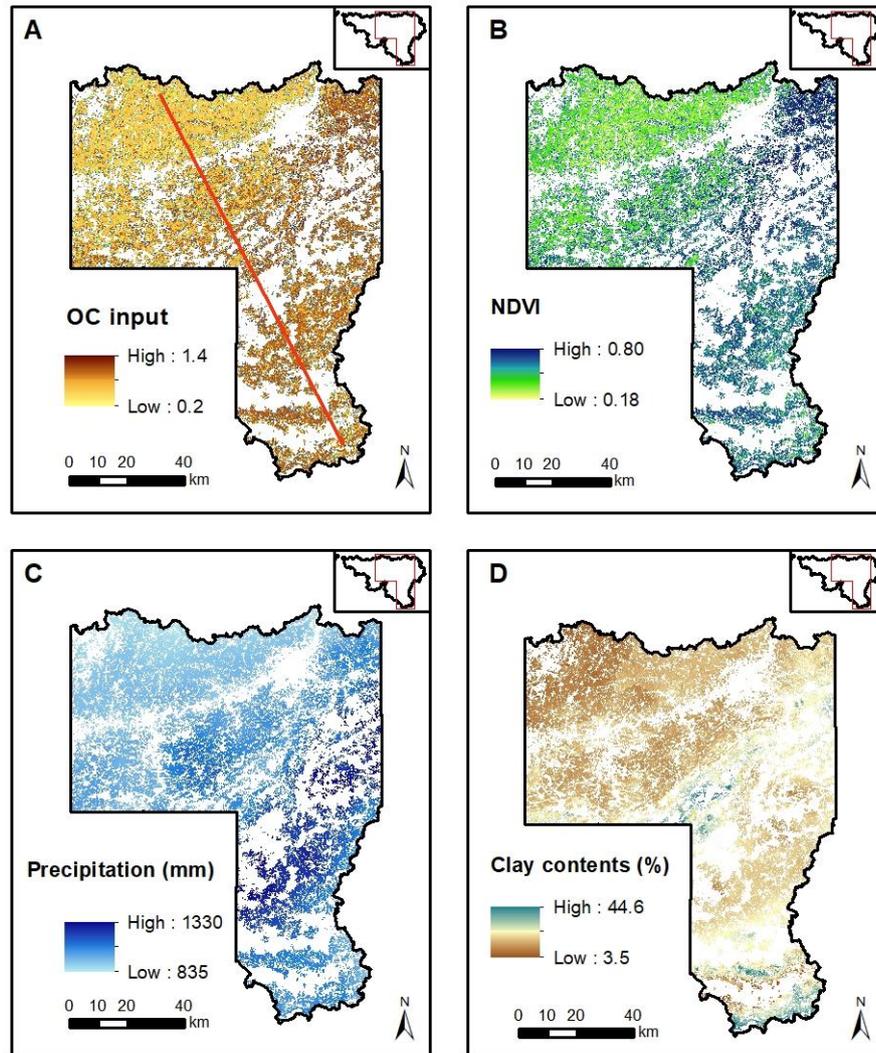


Fig. 2.4. Maps of Covariates (A)OC input (B)NDVI (C)Precipitation (D)Clay content. Red line: a transect selected for section 2.2.4

A further hyperparameters grid search was conducted using these five covariates to derive the best combination of parameters (shrinkage=0.05, depth of interaction=7, trees terminal nodes=8,

fraction of bagging=0.8, trees=500 (default)) and we constructed the final model. Following the completion of the modeling phase, the accuracy for the 10-fold cross-validation was obtained for the calibration dataset and then the model performance was evaluated using both the validation data and an external validation dataset.

2.2.3.3 Uncertainty Assessment (bootstrapping)

To measure and quantify the uncertainty of the predictions, we generated 50 GBM models based on non-parametric bootstrapping to calculate the prediction interval (Malone et al., 2011). Hence, we derived 50 predictions for each field. The mean value of the 50 maps were calculated as the final map, while the variance of the 50 maps as well as the mean squared error (MSE) estimated from the validation data was used to quantify the uncertainty (Malone et al., 2016). For each field, the 90% prediction interval is the square root of the overall prediction variance multiplied by the z-score value of 1.645 that corresponds to a 90% confidence level. (Eq. 2.6-2.10).

$$VAR_{all} = VAR_{boot} + MSE_{vali} \quad (2.6)$$

$$SE = \sqrt{VAR_{all}} * 1.645 \quad (2.7)$$

$$UPL = MEAN_{boot} + SE \quad (2.8)$$

$$LPL = MEAN_{boot} - SE \quad (2.9)$$

$$PIR = UPL - LPL \quad (2.10)$$

where VAR_{all} is the variance of the 50 times bootstrap predictions, MSE_{vali} is the mean square error estimated from the validation data (we use different MSE values for cropland and grassland), SE is the standard error, $MEAN_{boot}$ is the mean bootstrap prediction, UPL , LPL and PIR mean upper prediction limit, lower prediction limit and prediction limit range.

In addition, the Prediction Interval Coverage Probability (PICP, Malone et al., 2011) were calculated to validate the prediction interval.

The algorithm counts the percentage of the validating observed data within the respective prediction limits. Ideally the PICP is close to 0.9 for a 90% prediction interval, indicating that the uncertainty was correctly assessed.

2.2.4 Empirical mode decomposition

Empirical Mode Decomposition (EMD) is a method of signal decomposition which was introduced by Huang in 1998. The nonstationary and nonlinear signals are adaptively decomposed into intrinsic mode functions (IMFs) of different frequencies without setting of basic functions. Geospatial data are often nonstationary and nonlinear, which allows EMD to be applied to reveal attribute variation at different spatial scales. (Hu & Si, 2013; Zhou et al., 2021; Gong et al., 2020; Stallone et al., 2020) Within the study area, a transect (Fig. 2.4A red line) was created from northwest to southeast. This transect is 138.5km long and sampling points were created along the transect at 350m intervals. These sampling points were used to extract SOC contents from the final SOC map and covariate values from the covariate maps. Signal decomposition was performed on the SOC signals, then we calculated the correlation between SOC and environmental variables at different scales (i.e. IMFs).

2.3 Results

2.3.1 Descriptive statistics

The average SOC contents in the topsoil of the Walloon region were $17.45 \text{ g}\cdot\text{kg}^{-1}$, ranging from 6.68 to $63.61 \text{ g}\cdot\text{kg}^{-1}$ (Table 2.2). The calibration and validation datasets were comparable, with similar mean SOC contents and quantile values. Both datasets included extreme high and low SOC values, which can ensure the model is valid over a considerable range of SOC values. The external validation dataset showed a much lower mean and smaller range of SOC

contents compared with our sampled data because this dataset is almost entirely cropland data.

Table 2.2. Summary descriptive statistics of measured SOC content ($\text{g}\cdot\text{kg}^{-1}$) in calibration and validation datasets.

Dataset		All	Calibration	Validation	External validation
n		497	348	149	21,551
Min		6.68	6.98	6.68	3.00
Q25		10.05	9.94	10.34	11.00
Q50		11.79	11.66	11.95	13.00
Mean		17.45	16.96	18.15	13.85
Q75		17.14	16.44	19.99	15.00
Max		63.61	60.21	63.61	56.00
SD		12.33	12.08	12.49	4.78
Land use	Cropland	386	276	110	21,468
	Grassland	111	72	39	83

2.3.2 Model performance

The SOC predictions were satisfactory with most points close to the 1:1 line in a predicted/observed graph (Fig. 2.5), a R^2 value of 0.77, a RMSE of $6.14 \text{ g}\cdot\text{kg}^{-1}$, a RPD of 2.0 and a ME of $-0.20 \text{ g}\cdot\text{kg}^{-1}$, the ME value indicates a slight overestimation. One obvious problem of the results was overestimating predictions low values ($< 8 \text{ g}\cdot\text{kg}^{-1}$) and underestimating high values ($> 45 \text{ g}\cdot\text{kg}^{-1}$). It is clear that the standard deviation is small for low SOC values and large for higher values.

For the external validation dataset, the scatter plot between the observed and predicted SOC values as well as the histogram of each point's residue were plotted in Fig. 2.6. The R^2 value was much lower than the R^2 of the validation data, but this is reasonable given the narrower SOC distribution of the external validation dataset compared to the one used for model construction, as well as the different sampling protocols employed for validation (point scale) and external

validation (field scale). The small RMSE ($4.3 \text{ g}\cdot\text{kg}^{-1}$) also indicated a good model performance. The histogram showed an unbiased Gaussian-like distribution with an average value of $-0.2 \text{ g}\cdot\text{kg}^{-1}$. Given the size of the external dataset ($n=21,551$), the prediction was only slightly biased with a mean error of $-0.2 \pm 4.3 \text{ g}\cdot\text{kg}^{-1}$. 53% of the points have residues less than $2 \text{ g}\cdot\text{kg}^{-1}$, and 80% have residues less than $4 \text{ g}\cdot\text{kg}^{-1}$.

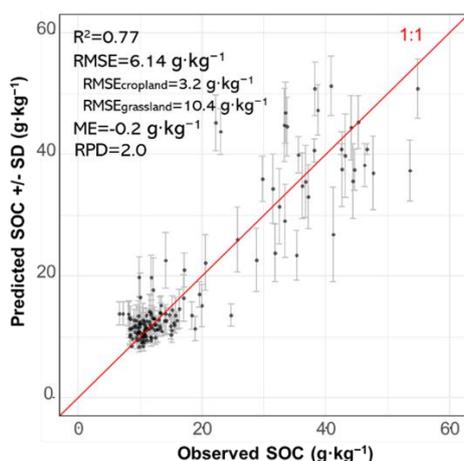


Fig. 2.5. Predicted vs. observed SOC contents of validation dataset ($n = 149$) (red line is the 1:1 line). The error bars are the standard deviations obtained by the bootstrapping.

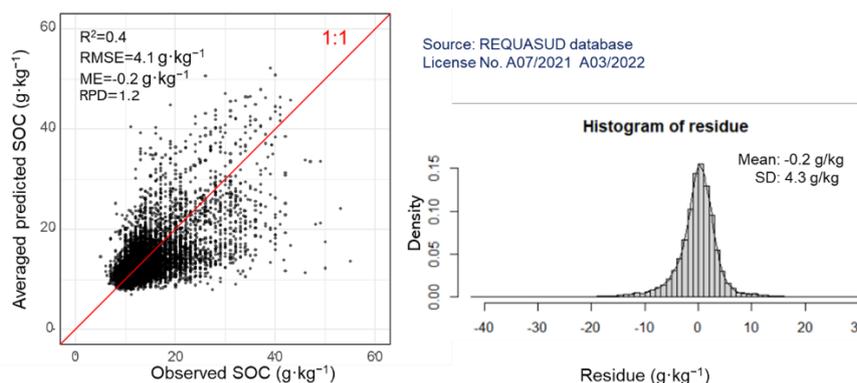


Fig. 2.6. Predicted vs. observed SOC contents of external validation dataset ($n = 21,551$) (red line is the 1:1 line) and the histogram of the residues.

The PICP value was 0.94 for cropland and 0.89 for grassland, close to the ideal value of 0.9, indicating that the prediction interval was a reasonable assessment of the prediction uncertainty. The values indicate a potential slight underestimation of prediction intervals for grassland and overestimation for cropland.

2.3.3 Importance of covariates

The GBM models generate the relative importance of each covariate during the modelling process, which is related to the degree to which the variation in a single covariate affects the accuracy of the prediction. Fig. 2.7 shows the importance of the covariates in 50 simulations in the form of raincloud plot. Among the five covariates, NDVI is the most important one, followed by OC_{input}, Precipitation, Clay content and DEM. However, there is some correlation between the factors, although it does not influence the model accuracy, it will cause confusion when assigning coefficients to a single tree learner, thus the interpretation of covariates importance loses its reliability.

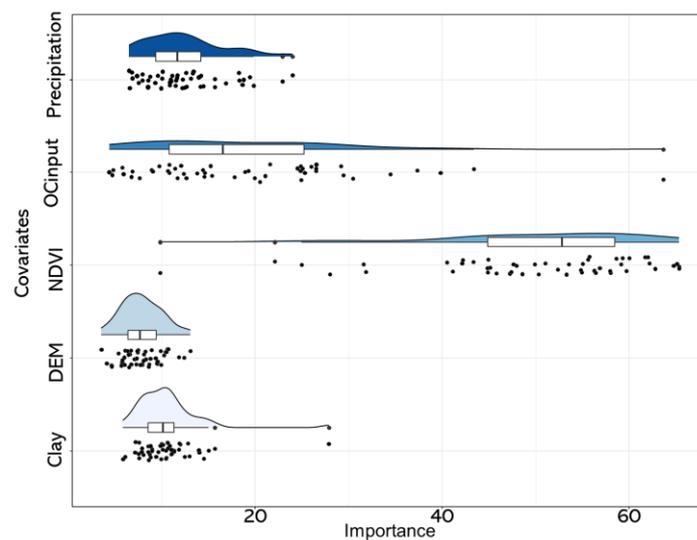


Fig. 2.7. The raincloud plot of the covariates' importance in 50 simulations, including an illustration of importance distribution (the cloud) with the raw data (the rain).

We further explored the relationship between covariates and SOC contents at different scales using the EMD approach (Fig. 2.8).

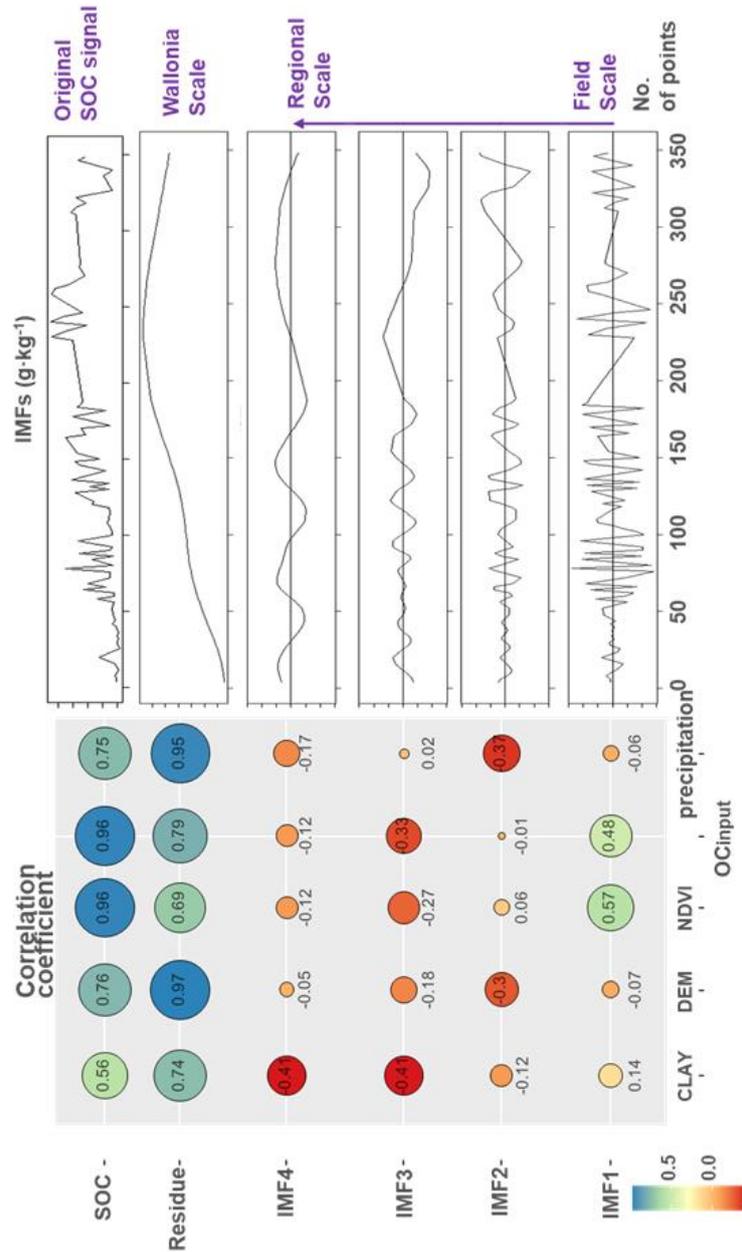


Fig. 2.8. Original SOC signal, residue and four IMFs of EMD, and correlation coefficients with covariates. (IMF: Intrinsic mode functions; EMD: Empirical mode decomposition)

The results of EMD for the SOC contents along the transect can be used as an additional explanation for the importance of the covariates. The original SOC signal with its residue and IMFS is displayed in the right part of Fig. 2.8 and correlation coefficients with covariates in the left part. First, from northwest to southeast, the SOC content shows an overall trend (residue part of EMD) of continuously increasing and then decreasing. This global trend showed a strong correlation with all factors. For Clay content, DEM and precipitation, the correlation coefficients with residue part were higher than those with the original SOC signal, suggesting that these three covariates influence the distribution of SOC at the scale of the entire region. After removing the influence of the global trend (IMF 4), NDVI and OC_{input} can capture the variability of SOC contents at field scale (around 0.7km) with a relationship value at 0.57 and 0.48.

2.3.4 Mapping SOC content and its uncertainty

Finally, we constructed a SOC map of a large part of the Walloon region at 30 m resolution (Fig. 2.9(A)). The regional map clearly shows the global distribution of SOC content, increasing first and then decreasing from northwest to southeast, with a mean value of $28.4 \text{ g}\cdot\text{kg}^{-1}$, a minimum of $6.6 \text{ g}\cdot\text{kg}^{-1}$ and a maximum of $60.3 \text{ g}\cdot\text{kg}^{-1}$. The SOC contents varied considerably and were distributed unevenly. The two river systems (Sambre-Meuse and Ourthe) divide the entire region into three gradients of SOC, with the lowest values to north of the Sambre-Meuse river system where croplands dominate, and the highest SOC to the south and east of the Ourthe river with mainly grasslands. We meanwhile mapped the spatial distribution of the 90% prediction limit range based on 50 times bootstrapping (Fig. 2.9(B)). Northern areas with a large number of soil samples have lower uncertainty and narrower prediction ranges, mostly less than $14 \text{ g}\cdot\text{kg}^{-1}$. In contrast, the south and east of Walloon region has few soil

observations, therefore prediction intervals are broader than in more densely sampled areas. Future sampling efforts should focus on the blue areas (mainly grassland) to minimize the existing uncertainty in surface SOC modelling.

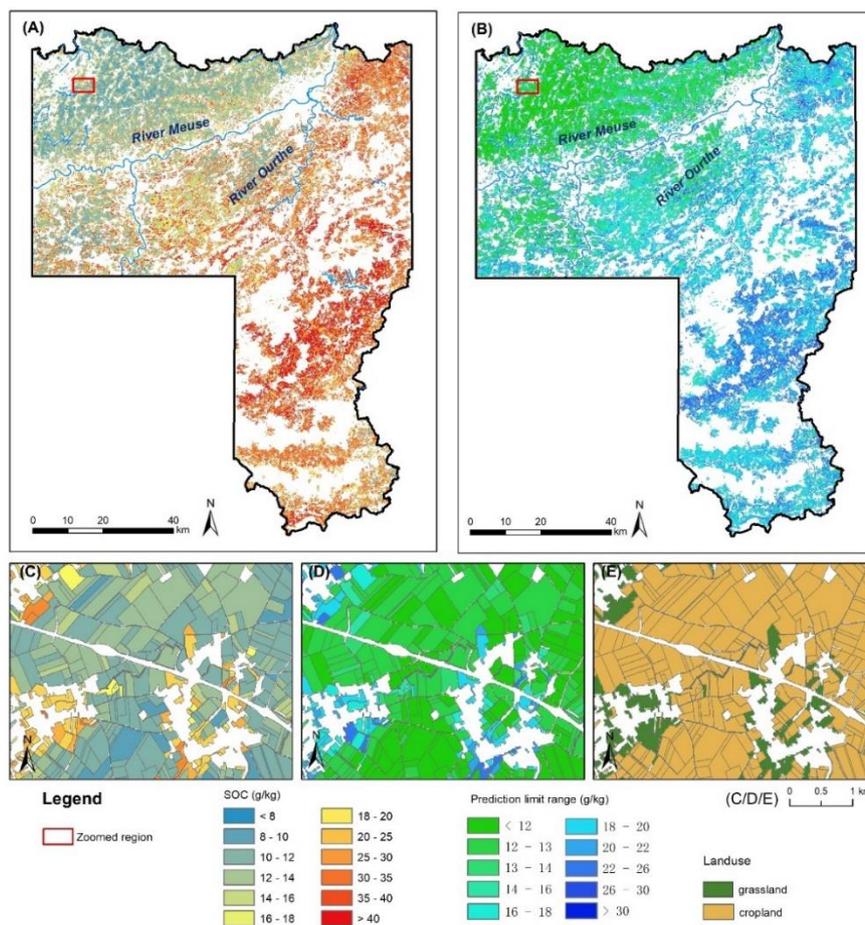


Fig. 2.9. Spatial distribution of (A) mean SOC content and relative 90% prediction limit range (B), zoomed to a smaller region (C) with its interquartile range (D), and land use map (E). The white areas are forests or built-up areas.

We also produced a detailed SOC map for an area north of the Sambre-Meuse river system (Fig. 2.9 (C)) together with its interquartile range (Fig. 2.9 (D)) and land use (Fig. 2.9 (E)). The distinction between fields is clear and for many fields SOC contents

and their corresponding prediction intervals are different from neighboring fields. In general, cropland fields have lower SOC values and smaller prediction limit ranges than grasslands.

2.4 Discussion

2.4.1 Quality of the Prediction

We compared our model to other published research and discovered that our prediction outperformed previous models applied in similar regions or at the similar scale. Meersmans et al. (2011) used multiple linear regression models to build the SOC concentration maps in the top 0.3 m of 1960 and 2006 for entire Belgium. For the 1960s legacy dataset, their R^2 was 0.42 and for the dataset of 2006 their R^2 reached 0.65. Chartin et al. (2017) obtained a reliable SOC stock map (2004-2014) for the Walloon region based on Generalized Additive Model with an R^2 of 0.64.

Many other regional SOC maps have become available at high resolution (10-30 m) in recent years, triggered by free high-resolution remote sensing data. In order to assess Sentinel-2's ability to predict topsoil properties, Vaudour et al. (2019) chose two contrasting pedoclimatic environments in France: one temperate region (Versailles Plain) and one Mediterranean region (Pène Valley) and built SOC models based on partial least squares regressions with R^2 of 0.56 and 0.02 respectively. Zhou et al. (2020) employed DEM derivatives, Sentinel-1, and Sentinel-2 data to map SOC content for the southern part of Central Europe (encompassing Slovenia and a small portion of Austria and Italy), and the boosted regression trees beat other machine learning approaches with an R^2 of 0.57.

Hence, we successfully predicted SOC content for the Walloon region with a GBM model and a series of covariates and obtained a

validation R^2 value (0.77) that was slightly higher than the one for the studies mentioned above (R^2 range from 0.02 to 0.7). Furthermore, the SOC map for the Walloon region not only provides a higher spatial resolution (field scale) but is also more recent (2015-2019). Thus, we confirm the effectiveness of our entire workflow and model procedure for large-scale and high-resolution digital soil mapping.

In addition, the better performance for croplands (RMSE=3.2 $\text{g}\cdot\text{kg}^{-1}$) than grassland (RMSE=10.4 $\text{g}\cdot\text{kg}^{-1}$) can be explained because of the more detailed information on OC input. After all, for croplands, data such as the crops grown each year, an estimate of their biomass input to the soil and whether a cover crop was cultivated are available. For grasslands there is no quantitative data on biomass input, composition of the grazing intensity or number of cuts when mown.

2.4.2 SOC map products and its possible opportunities

2.4.2.1 Zoomed into the field scale

One of the key accomplishments is the ability to continuously zoom in on the SOC map, down to the field scale and clearly demonstrate the difference in SOC content (Fig. 2.9). We compared our maps to published open-access SOC maps of the Walloon region (<https://geoportail.wallonie.be/>), by zooming in to agricultural landscapes north and south of the Sambre-Meuse river system (Fig. 2.10). The Walloon SOC map (A2/B2) followed the same broad patterns as our map (A1/B1), with small patches of high SOC values in the grasslands. The Walloon SOC map shows an overall trend of SOC distribution, for example, in A2, there are low values in the northern part and high values in the southern part. Evidently, it cannot give a good indication of the differences between fields when compared with the field-scale map.

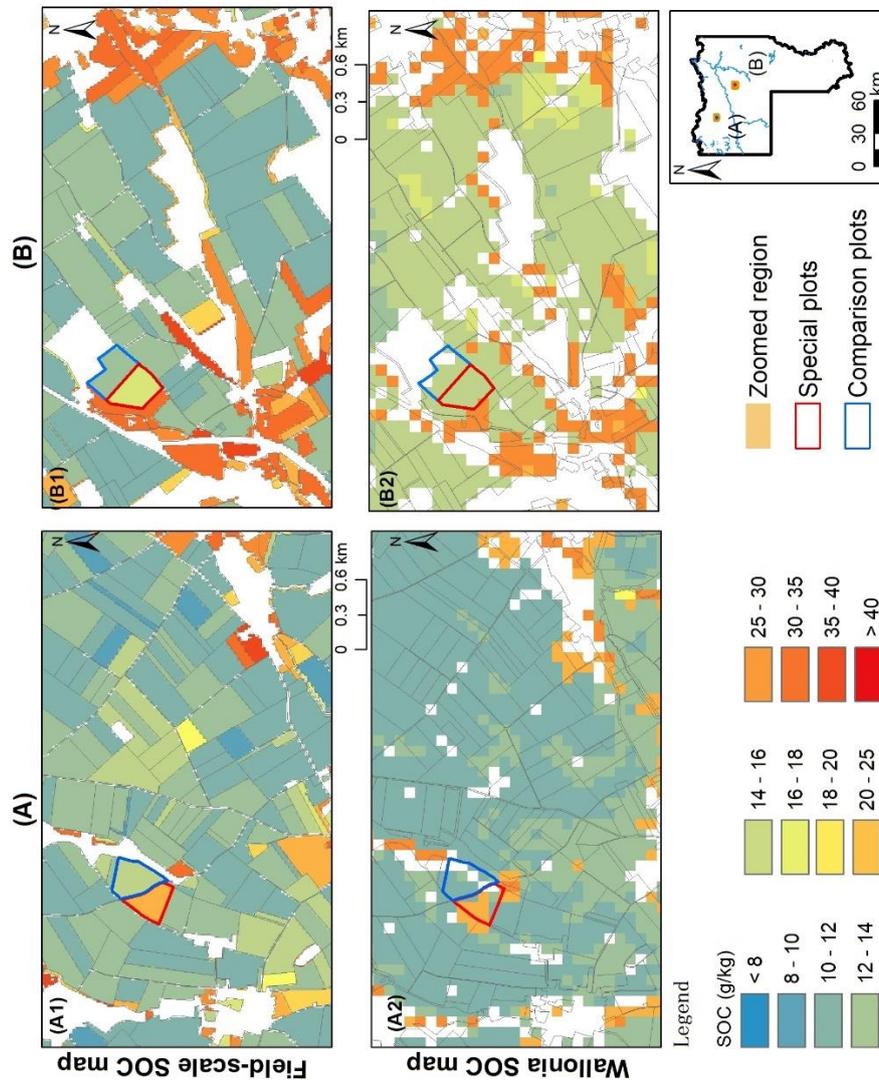


Fig. 2.10. Comparison of zoomed SOC content maps from our work (A1) (B1) and the Walloon SOC map (A2) (B2).

After all, no specific covariates were employed in the modelling procedure to represent the difference across the fields. Moreover, the coarse pixels do not always represent a single field. As a result, The Walloon SOC map, designed for evaluating the state of the soils in the entire Walloon region, is less specific in details and not sufficient to characterize the spatial distribution pattern of SOC at the field scale.

Our work will not only address the gap by providing a credible SOC map for the Walloon region but will also allow SOC maps to finely characterize the field scale. This approach provides a baseline map both for farmers making management decisions about their fields, and for authorities evaluating the state of the environment.

We selected two typical fields which were characterized by slightly higher SOC contents than the surrounding fields, whereupon we listed the crop types of these plots (red border) and the neighboring plots (blue border) over the last five years (Fig. 2.10 and Table 2.3).

Table 2.3. Crop types and total OC input (include CC) in the field over a five-year period. (CC indicates the existence of a cover crop after the main crop)

Zone	Type	2015	2016	2017
A	typical	Maize silage	Winter wheat	grassland
	Contrast	Potato	Winter wheat	Peas
B	typical	Winter wheat (CC)	Flax	Winter wheat
	Contrast	Winter wheat	Potato	Winter wheat (CC)
Zone	Type	2018	2019	OC input (kg·C·m ⁻²)
A	typical	grassland	grassland	3.77
	Contrast	Sugar beet	beans	2.23
B	typical	Winter rapeseed	Winter wheat	2.96
	Contrast	Flax	Sugar beet	2.70

The field with high SOC content in zone A was converted from cropland to grassland in 2017, while the neighboring field remained under cropland. This in agreement with Lugato et al. (2013) who used the CENTURY model to simulate the carbon sequestration potential of different management practices on 12%-28% of European arable land, and the results showed that the change in land use from arable to grassland had the highest carbon sequestration potential with respect

to all the others simulated practices. Poepflau & Don (2013) also stated that conversion from cropland to grassland resulted in an average SOC accumulation of $17 \pm 5 \text{ Mg ha}^{-1}$ SOC in the topsoil (0-30 cm).

For the field in Zone B, the high SOC value is due to the higher OC input over the five-year period. The typical field had four crops (cereals and rapeseed) with a high OC input, while the neighboring field had only two high OC input crops (cereals). It is generally well-known that increase in C inputs is the most efficient strategy to enhance SOC stocks (Fujisaki et al., 2018; Martin et al., 2021). In addition, FAO pointed out that rotation of at least three different crops (including repetitive wheat, maize or rice) will alter pest cycle, diversify rooting patterns and rooting depth and finally increase humus formation and reduce risk of pest and weed infestations (Palm et al., 2014; FAO, 2011; West & Post, 2002). One fundamental principle of conservation agriculture is keeping the soil covered, when the gap between harvesting one crop and establishing the next is too long, cover crop would be cultivated. Cover crops aid in the accumulation of organic matter in the surface soil horizon by reducing water losses, weed infestation and promoting biological soil tillage through their rooting (Alvear et al. 2005; Hobbs et al., 2007).

These two fields adequately reflect the positive impact on SOC contents due to different management practices. It can not only serve as a reference for agricultural stakeholders in achieving their multiple goals (e.g. improve soil productivity, increase soil carbon sequestration, mitigate greenhouse gases), but also offer the possibility of monitoring the implementation of conservation agriculture policies.

In addition, within some zoomed-in areas, we attempted to investigate the relationship between land use history and SOC contents of the fields (Fig. 2.11). Copernicus CORINE Land Cover data (100 m) for the study area were downloaded from GEE for a total

of five periods (1990, 2000, 2012, 2016 and 2018). Some fields changed from forest to agricultural land between 1990 and 2000 (Fig. 2.11. red plots). We summarized the SOC contents of these fields and their neighboring paired fields. The SOC mean value of these fields cropped for more than 30 years (blue fields) were $16.9 \text{ g}\cdot\text{kg}^{-1}$, which was obviously smaller than the recently deforested fields (red fields) of $24.3 \text{ g}\cdot\text{kg}^{-1}$. This result demonstrated the declining SOC trend following the conversion from forest to cropland. Villarino et al. (2017), Berihu et al. (2017) and Osinaga et al. (2018) also confirmed this perspective in their study. According to Villarino et al. (2017), 10 years of cropping after deforestation will bring roughly 30% SOC loss at 0–30 cm depth. It's worth noting that carbon losses are not limited to the topsoil, but also affect the vertical SOC distribution and deeper SOC stock, ultimately contributing to the greenhouse gas inventories.

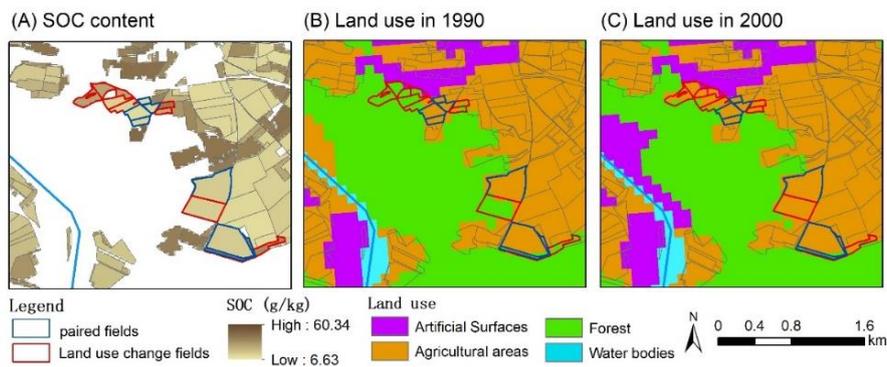


Fig. 2.11. The distribution of SOC content ($\text{g}\cdot\text{kg}^{-1}$) (A) in an area with recent land use change (B): land use in 1990, (C): land use in 2000.

2.4.2.2 Application example of SOC products for Belgium: derived soil aggregate stability class map

In addition to serving as a baseline map to provide farmers and agencies with SOC information on agricultural land, the SOC map as well as the bootstrapped predictions can be further calculated to have additional applications.

We compiled an aggregate stability class map (Fig. 2.12) from the final SOC map and the threshold given by Shi et al. (2019).

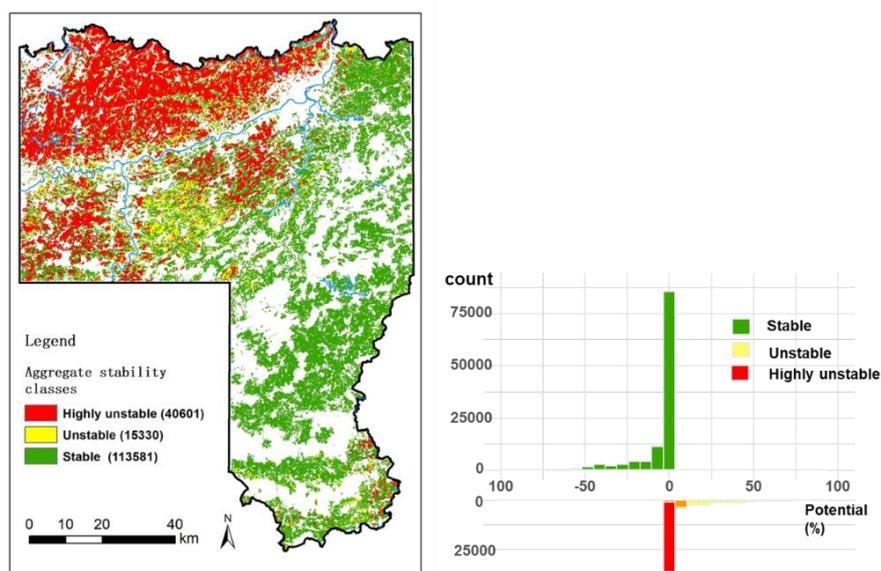


Fig. 2.12. Map of different aggregate stability classes (i.e. highly unstable, unstable and stable) and histogram of misclassification potential (The probability of not belonging to the current class among the 50 predicted results, negative values indicate downgrading).

Based on the assessment and analysis of soil aggregate stability in the Belgian Loam Belt, they raised two critical SOC thresholds (i.e. 15 $\text{g}\cdot\text{kg}^{-1}$ and 20 $\text{g}\cdot\text{kg}^{-1}$) to separate the aggregate stability into “Stable”, “Unstable” and “Highly unstable” classes. Considering the prediction errors, potential misclassification risk was also calculated based on 50 times bootstrapping predictions (percentage of prediction that over/under the 20 $\text{g}\cdot\text{kg}^{-1}$ limit for (Unstable/ Highly unstable)/ Stable class).

In 169,512 agricultural fields of our study, 67% of them are "Stable," 69% of these stable fields are entirely stable (all 50 times predictions were bigger 20 $\text{g}\cdot\text{kg}^{-1}$, potential=0), and mostly distributed in the eastern and southern regions. In addition, 9% of the plots were

“Unstable”, with 24% being “Highly unstable”, and the probability that these fields are misclassified stable fields is also extremely small (only 0.4% has potential more than 50%). These fields were mainly in the northwest region, which is a hotbed of intensive agricultural practice. These soils have low physical stability and are more susceptible to surface crusting, soil erosion, and soil degradation (Loveland & Webb, 2019). Furthermore, crop yields in agricultural land may decrease due to the soil's reduced capacity to recycle nutrients, which may lead to a decrease in crop residue's SOC input into the soil. A vicious cycle is set in motion. It demonstrates the critical importance of maintaining appropriate levels of SOC, which can be achieved through fallowing, conversion to grassland or forest, and reasonable organic inputs (Abiven et al., 2009).

2.4.2.3 Uncertainty assessment and its monitoring implications

The assessment and communication of uncertainty is an open challenge (Arrouays et al., 2020; Chen et al., 2022). Uncertainty should be able to provide reliable information and guidance to policy makers and other stakeholders, not just scientists and modelers (Poggio et al., 2021). It's also debatable if the *GlobalSoilMap* specifications' requirement of 90% prediction intervals (PIs) is necessary (Chen et al., 2022). Andries et al., (2021) conducted a series of semi-structured interviews with experts in soil health and policy, nearly 70% respondents expressed PICP should be greater than 90%, and 15% respondents considered 70% is acceptable. While according to Helmick et al. (2014), the actual width of 90% intervals may be bigger than the need of practical application. A PIs of 75% may be sufficient to support a decision.

Furthermore, a single PIs value appears to be of limited utility for expressing and understanding uncertainty. The communication of uncertainty will be more effective when the consequences of

uncertainty (e.g., subsequent sampling design, change monitoring) are delivered as a final product (Arrouays et al., 2020; Chen et al., 2022). It's worth noting that change monitoring on a temporal scale is difficult for regional SOC content with relatively high RMSE and wide PIs in each pixel (Kempen et al., 2018). At the level of resolution of the grid, the uncertainty is big, for example, we calculated 90% PIs, adding or subtracting the limit range to the current SOC predictions will result in an unreasonable SOC level, thus also lead to the meaningless of calculating the time when the SOC changes can be detected based on this. Monitoring over time may be achieved by using narrower PIs or aggregating to higher geographical levels. Kempen et al. (2018) clarified that predictions could be aggregated at the level of land cover classes, geographic regions, or climatic regions, which will reduce the uncertainty associated with the predictions and allow for the detection of changes over time.

2.4.3 Controlling factors of SOC patterns

We used EMD to classify the effects of environmental covariates on SOC distribution at two scale ranges, Walloon region and field-scale controls, allowing us to better understand the role of various covariates.

At the scale of the Walloon region, climate, topography, soil texture and biomass shaped the overall pattern of SOC contents' spatial distribution. In agreement with the literature, SOC contents show a significant positive correlation with elevation, rainfall, and clay content. The increase in altitude brings a decrease in temperature, which slows down the activity of microorganisms in the soil and the rate of mineralization as well, resulting in a further accumulation of total SOC (McBratney et al., 2003; Tashi et al., 2016; Ali et al., 2017). Precipitation determines net primary productivity and thereby affects the litter carbon pool and soil carbon pool (Wiesmeier et al., 2019;

Giardina & Ryan, 2000; Follett et al., 2012). However, rather than precipitation alone, water balance (and its proxy, the ratio of precipitation to evapotranspiration) plays a more critical role in regulating decomposition rates (Yang et al., 2002). For example, in Chapter 4, both precipitation and open pan evaporation jointly influence soil carbon decomposition rate. This factor is often overlooked in commonly used DSM covariates and should be integrated into future work. With increasing clay content, a decrease in decomposition rates was reported (Giardina et al., 2001).

At the field scale, the results show that OC_{input} and NDVI play a dominant role in the distribution of SOC, with an influence distance of 0.7 km. Some SOC models, in particular the ones that characterize SOC mineralization with first-order kinetics, imply that steady-state SOC stocks are proportional to C inputs (Paustian et al., 1997; Martin et al., 2021). Chenu et al., (2019) suggested that management practices influence SOC stocks by increasing SOC inputs (e.g. crop rotations, agroforestry, cover crops, crop residue return) to soil and decreasing SOC losses (e.g. no tillage, water management, decreasing erosion rates). We quantified the factors of carbon input at the field scale but unfortunately could not address the factors related to SOC losses. In addition, Stevens et al. (2015) decomposed variability of soil properties by spatial scales using filtered kriging and the results obtained agreed with our work, i.e. the environmental factors related to field/farm management control the SOC variability at short range, equals 700m. Meanwhile, we obtained some information on the structural characteristics of agricultural holdings via Eurostat (<https://ec.europa.eu/>). As an agricultural census, The European Union started Farm structure survey (FSS) from 1966. In FSS 2013, the averaged utilized agricultural area of each holding in Walloon region was 54.9 ha. Assuming that the management area is a square, the side length would be 740m, which is also compatible with the spheres of

influence of OCinput and NDVI obtained above. This result is reasonable since each holding would generally adopt the same management practices within its jurisdiction.

2.4.4 Limitations and future perspectives

First, the fundamental issue is a scarcity of well-distributed soil samples within the study area. The point observations exhibited distinctly “cluster” zones in northern Walloon region, where the cropland fields are often under intensive management, thus the dataset and model developed are more represented by fields under intensive practices and have poorer prediction performance for fields in the south. Also, the results of standard cross-validation will be heavily influenced by the performance in high sampling density areas and resulting in an overly optimistic view of model accuracy (Poggio et al., 2021; Meyer et al., 2018). Hence, proper sampling strategy (Minasny & McBratney 2006), declustering of the data (Deutsch and Journel, 1998) or spatial cross-validation (Ploton et al., 2020; Meyer et al., 2018) should be explored further.

Second, another major limitation of this study is directly using the crop type data provided by the agricultural census and OC input data from the literature. However, in practice, the accuracy of recording crop types is not fully guaranteed, and a midway shift in harvest practices may not always be recorded in detail. For example, if a grain maize was originally planted, which is usually harvested in November, but the farmer may harvest the maize as maize silage earlier based on yield estimates or a change in actual demand. The difference is not particularly significant for the farmer, but the difference in carbon input between these two harvesting methods is significant (almost twice; Table A.1). One way to address this problem is to check crop residues after harvest based on remote sensing images using indicators such as The Normalized Burn Ratio

(NBR2) index (Demattê et al., 2018; Dvorakova et al., 2020). Another solution is to conduct crop classification for each field with long time series satellite images (Piedelobo et al., 2019; Kussul et al., 2017).

Finally, not all management practices have been considered. On one hand, the amount of manure and compost applied to the fields is lacking, which is an important source of organic carbon input. On the other hand, minimum mechanical soil disturbance is one of the main principles of conservation agriculture which was not mentioned earlier and considered in this work either. Less intensive tillage practices and no tillage can mitigate negative impacts on soil quality and preserve SOC (West & Post, 2002; Haddaway et al., 2017). So, is it possible to identify tillage or no tillage at field-scale with satellite images? Examples show that if ploughing occurs immediately after the harvest, the NDVI will show another small drop based on the original lowest point. With the help of SAR backscatter data of Sentinel-1 and optical NDVI data of Sentinel-2 (Satalino et al., 2018), tillage detection is feasible in further research and will be practically useful.

2.5 Conclusion

We characterized spatial patterns in soil organic carbon content and the accompanying modeling uncertainties for part of the Walloon region. We comprehensively summarized the modelling results and analyzed the covariates at different scales:

1. Using the GBM model, a reasonable result was obtained, which can explain 77% of the variance in the SOC contents for Wallonia, but there is still obstacle gap between the " validation accuracy" and policy requirements.
2. Farm-scale maps can make up the gaps in existing DSM work, not only by providing more accurate SOC contents for each field, but also by analyzing and documenting management practices. Proper

crop rotation, conversion to grassland and sowing of cover crops all contribute to organic carbon accumulation.

3. Climate, topography and soil texture factors shape the distribution of SOC contents across the Walloon region, the differences between fields are more likely to be generated by differences in organic carbon inputs from crop residues.

The results of this study are able to improve understanding the controlling factors for SOC contents at different scales and serve as a baseline field-scale map to monitor SOC dynamics, formulate land management strategies and support economical- agricultural policy.

Chapter 3.

Mapping conservation agriculture fields and management practices

The content of this chapter was submitted as Zhou, Y., Ferdinand, M., van Wesemael, J., Dvorakova, K., Baret, P., Van Oost, K., van Wesemael, B., 2024. A framework for mapping conservation agricultural fields using optical and radar time series imagery (Under review)

3.0 Outline

The importance of conservation agriculture (CA) is undeniable, both for improving soil health and offering a viable path towards achieving carbon neutrality. However, to date, survey statistics on the extent of conservation agriculture were based on farmer declarations or field inspections. This is a major impediment to the promotion or monitoring of conservation agriculture. Here, we collected the management practices of a total of 247 fields under conservation agriculture in the Walloon region of Belgium in 2020-2021, with the aim of developing a classification model for the prediction of conservation agriculture by combining remotely sensed data with census data. We identified seven variables in the model, linked to each of the three main principles of conservation agriculture (crop diversification, maximum soil cover and minimum mechanical soil disturbance). The number of different annual crops and cereals in the rotation was obtained from the agricultural census. For the extent of

soil cover, the Google Earth Engine (GEE) platform was used to obtain a time series of optical remote sensing images (2015-2020, Sentinel-2, Landsat-7, Landsat-8) and precipitation data. We then analyzed the variation of spectral indices such as the Normalized Difference Vegetation Index (NDVI) and the Normalized Burn Ratio (NBR2) and constructed indicators to distinguish between bare soil and cover crop. For minimum mechanical soil disturbance, in addition to the above data, radar data (Sentinel-1) were also obtained from the GEE platform to establish a tillage practice model. Subsequently, the Random Forest (RF) classification method was used to construct a classification model distinguishing fields under conservation from those under conventional practices. The results of a ten-fold cross-validation showed a good overall accuracy of 92 %. The model was utilized to classify the farming systems in all croplands of the Hesbaye region of Belgium. The results show that 15.5% (2,875 fields) out of 18,516 cropland fields can be classified as conservation agriculture. These fields tend to adopt non-inversion tillage and have diverse crop rotations.

3.1 Introduction

Agricultural soils are suffering from increasing threats such as erosion, declining organic matter and biodiversity, pollution, sealing, compaction and salinization (Ferreira et al., 2022; Panagos et al., 2018; Foley et al., 2011). To address the challenges of soil degradation and promote sustainable agriculture, the Food and Agriculture Organization (FAO) advocates for the adoption of conservation agriculture (CA) as a set of widely adopted agricultural production systems that improve or maintain crop production while restoring degraded land and providing sustainable production systems. The FAO (fao.org/conservation-agriculture/) defines CA as a set of practices based on three main principles: (1) crop species diversification; (2) maximum soil organic

cover (including living mulch (cover crops, annual crops, temporary grasslands) and dead mulch (crop residues, decaying leaves, bark, manure)) and (3) minimal mechanical soil disturbance. Adopting these practices will have numerous benefits for soil health (Page et al., 2020; Rodríguez et al., 2022), such as increasing aggregate stability and soil quality (Castellanos-Navarrete et al., 2012; Sithole et al., 2019; Jat et al., 2019), reducing soil erosion (Petito et al., 2022), improving soil biological abundance and diversity (Ayuke et al., 2019; Giraldo-Perez et al., 2021), and increasing crop yields and farmers' incomes (Naab et al., 2017; Pradhan et al., 2018). However, current methods of distinguishing fields under CA from those under conventional agriculture, or traditional agriculture (TA), are still limited to direct observation in the field in combination with farmers' interviews. This is labor-intensive and time-consuming, making it difficult for large-scale inventories of CA applications. Remote sensing techniques have proven to be an efficient and effective tool for generating various types of agricultural data (crop types, crop yields, management technique, critical soil properties such as soil organic carbon (SOC) content) (Griffiths et al., 2019; Sparks et al., 2022; Blickensdörfer et al., 2022; Evans & Shen, 2021; Zheng et al., 2014; Castaldi et al., 2019). However, the identification of fields under CA remains challenging.

The breakthrough in addressing this challenge is that remote sensing techniques can be used to effectively predict or assess each principle of CA separately. By integrating the indicators for all three principles, a comprehensive assessment of CA can be achieved. First, to measure species diversification, it is crucial to identify the crop planted each year. This typically requires field-scale agricultural census databases, such as the Land Parcel Information System (LPIS) used by Zhou et al. (2022) for the Belgian region. However, such field-scale data are often scarce. An alternative approach is to use remote sensing data for crop type mapping (Johnson, 2019; Belgiu &

Csillik; 2018). Since crop types determine the entire phenological cycle from seeding to harvest within a relatively short time (Blickensdörfer et al., 2022), the spectral trajectories of different phenological stages can be captured in detail using high-resolution, long-term remote sensing data (Preidl et al., 2020). Second, to determine winter cover crops, that are not reported in the agricultural census, it is necessary to differentiate remote sensing signals from soil, main crops and cover crops (Zhou et al., 2022). Researchers can distinguish between bare soil and vegetation using vegetation indices and soil indices (Mzid et al., 2022; Dvorakova et al., 2023). Subsequently, by integrating phenological information with known main crop types, and separating the main crops and soil signals, if clear plant growth signals are detected during the time window where cover crops might be present, cover crops can be determined using vegetation index thresholds (Zhou et al., 2022; KC et al., 2021). Finally, time-series of satellite imagery have been proven to be an effective tool for monitoring and evaluating tillage practices (Sonmez & Slater, 2016; Zhou et al., 2021; Liu et al., 2022). Inversion tillage (i.e. inverting topsoil layers, usually to a depth of 30 cm in Belgium) buries most of the crop residue in the soil, whereas non-inversion tillage retains most of the crop residue in the topsoil. This distinction in tillage practices can be detected through the differences in residue cover. The cellulose and lignin in crop residues and soil will exhibit distinctive absorption characteristics in the shortwave infrared (SWIR) range, particularly around 2100 and 2300 nm (Elvidge, 1990). The Cellulose Absorption Index (CAI) has been shown to accurately represent these features (Bai et al., 2021; Hively et al., 2021; Pancorbo et al., 2023), but it requires high spectral resolution in the SWIR range (centered at 2000, 2100, and 2200 nm; Nagler et al., 2000), which is not available on commonly used satellite platforms such as Landsat and Sentinel (Dennison et al., 2023). As an alternative, broadband

indices like the Normalized Difference Tillage Index (NDTI, van Deventer et al., 1997), also known as the Normalized Burn Ratio 2 (NBR2), have been adopted. However, while NDTI is useful, it can be affected by interference from green vegetation and moisture content (Lamb et al., 2022; Wang et al., 2013), which may introduce uncertainties into the assessment. Quemada et al. (2018) performed moisture correction on irrigated fields before using a tillage index for crop residue estimation. In addition, soil surface roughness can serve as another indicator for detecting tillage, but its changes are more ephemeral, making it difficult to detect accurately with satellites passing every 5-10 days. Radar data has been proved for estimating bare soil surface roughness in many studies (Zribi & Dechambre, 2003; Baghdadi et al., 2002, 2018). Azzari et al. (2019) found that combining optical Landsat satellite data with radar-based Sentinel-1 data could improve classification accuracy for tillage practices in the North Central US Region.

Here, we integrate and improve existing methods to develop a classification model for CA. To achieve this, time series of optical data from Landsat-8, Landsat-7, Sentinel-2 and radar data from Sentinel-1 combined with environmental data are good sources of data support. In addition, the Google Earth Engine (GEE) platform can store large volumes of remote sensing data in a cloud environment, allowing us to save time in satellite data processing and enabling efficient, fast, easily replicable and scalable workflow.

The specific objectives can be summarized as follows:

1. To extract quantitative indicators from time series of optical and radar imagery to estimate soil cover and tillage practices.
2. To develop a classification model to distinguish CA from TA using indicators constructed based on the three principles of CA.
3. To map the distribution of fields under CA in Hesbaye region of Belgium using the classification model.

4. To explore the contribution of covariates to the classification results of individual fields.

3.2 Method and materials

3.2.1 Study area

The study area is located in the southern part of Belgium, in the so-called Walloon region (Wallonia) (Fig. 3.1). Wallonia is a landlocked region covering a total area of 16,901 km². The altitude progressively increases from 100 m above sea level in the north-west to a maximum of 700 m in the south-east, leading to an increase in annual total precipitation from 800 to 1400 mm, and a decrease in the annual average temperature from 10 to 8 °C. Changes from deep sandy loam and silty soils to shallow silty and stony soils occur in the same direction, along with a shift from intensive arable agriculture to more extensive cattle grazing (Chartin et al., 2017; Goidts & van Wesemael, 2007).

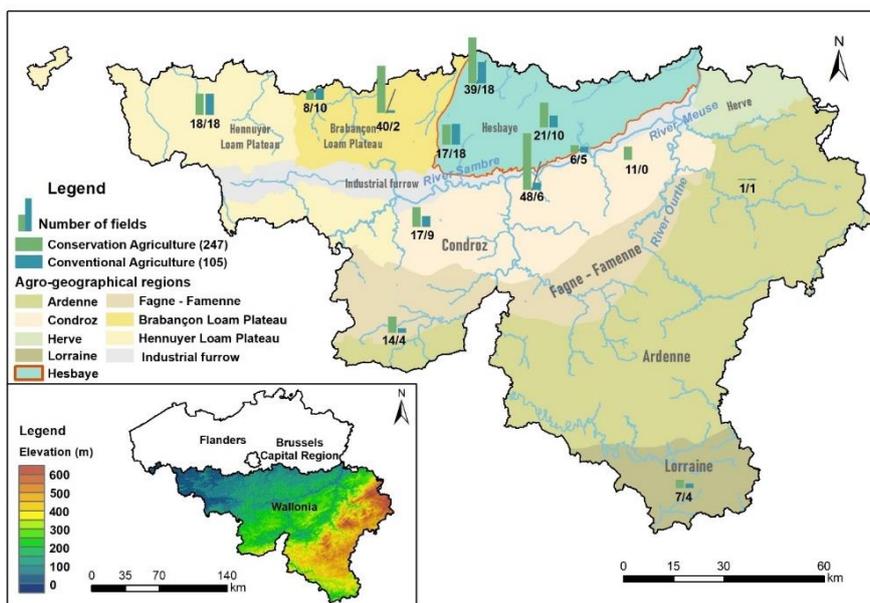


Fig. 3.1. Digital elevation model and agro-geographical regions of the Walloon region, the bar plots display the number of conservation (green bar) and conventional (blue bar) agriculture fields in the survey.

There are nine agro-geographical regions in the Walloon territory (Fig. 3.1& section 1.9). The Hesbaye region has fertile silt loam soils with major crops such as cereals, sugar beet, potatoes and flax. Additionally, the Hesbaye region is entirely located within the Nitrate Vulnerable Zones (NVZ), which as designated by the European Union's Nitrate Directive. This regulation mandates that farmers in the NVZ region are obliged to cover the soil with cover crops during the winter months to reduce nitrate leaching and protect water quality. Field management data was collected across all Walloon regions, while the conservation agriculture mapping was done only in the Hesbaye region.

3.2.2 Field data

3.4.2.1 Data on fields under conservation agriculture

A preliminary list of farmers practicing CA was acquired involving twelve public and private institutions, eight farmer associations, two researchers from Belgian universities, and seventy farmers and/or agricultural entrepreneurs through the social network Facebook. We conducted semi-structured telephone interviews and cross-referencing with these farmers to verify and gather the information on Wallonia's CA fields (Ferdinand & Baret, 2024). The year of implementing CA practices and details on the tillage were two of the main questions asked during the interviews. At the end of the interviews, the respondents were requested to furnish a list of additional Walloon farmers who they knew to follow CA practices. This kind of “snowballing” strategy assisted us to obtain a more diverse source of data and more effectively target the next interviewee.

A total of 221 Wallonia farmers were contacted, and the farmers whose fields did not meet the CA criteria were removed from the database. The main prerequisites were: (1) at least five years of CA practice. (2) Implementing conservation tillage, which included direct seeding (i.e. no-tillage, growing crops without any soil preparation),

non-inversion tillage (a soil preparation practice involves fragmentation, mixing and burial, without horizon inversion) and occasional inversion tillage (horizon inversion with reduced tillage frequency or depth less than 15 cm) (Ferdinand & Baret, 2024). Following the digitization of the interviews, information on each farmer covered the type of tillage, depth of tillage, direct seeding or not as well as livestock presence or absence.

Between November 2021 and February 2022, 328 fields under CA practice were identified through semi-structured interviews with 28 farmers dispersed throughout eight of the nine agricultural regions (all but the Herve region). These fields are mainly concentrated in the Hesbaye, Brabanton Loam Plateau and Hennuyer Loam Plateau area in the north of the Walloon region (Fig. 3.1).

3.4.2.2 Data on fields under conventional agriculture

In the course of the semi-structured interviews, we also asked the farmers about the fields adjacent to their own fields. In particular, whether or not the surrounding fields were tilled at regular intervals and to a depth of more than 20 cm. If so, it indicated the use of inversion tillage, then we recorded these fields as TA fields. A total of 117 fields were obtained by this method. Furthermore, during the subsequent processing (see section 2.4), due to the requirement for high density of time series data, the fields that did not meet the requirement were removed. We only focus on CA in croplands and therefore exclude all temporary and permanent grasslands. In the end, 247 fields under CA and 105 fields under TA were retained.

3.4.2.3 Agricultural census data

The annual crops and the field lay-out are collected from the website of the Walloon Region (<https://geoportail.wallonie.be/>), which offers free public access to basic geographic dataset and maps on the topography, administration, environment, transportation, and tourism.

We have requested the data from the Land Parcel Information System (LPIS) for the years 2015 to 2020, with each data file containing individual field boundaries and annual crop type for all fields for which subsidies under the Common Agricultural Policy (CAP) program are requested, stored as polygons. Based on farmer declarations, field verification, and remote sensing checks, crop information and field boundaries are updated annually.

The Hesbaye region contains 44,893 parcels of agricultural land, totaling 124,058 ha. First, we performed a preliminary screening, keeping only those fields with an area larger than 1 ha. Then, only those fields were kept that had not been recorded as temporary or permanent grassland throughout the previous six years. Finally, 18,740 cropland fields covering 98,289 ha are available.

3.4.2.4 External validation dataset

Soil Capital© provided the external validation accuracy matrix based on our output maps of cover crop and tillage practice. Soil capital is an independent agronomy company founded in 2013, currently focused on developing a carbon payment program, which has already been implemented in Belgium, France, and the United Kingdom, in order to accelerate the transition to regenerative agriculture.

For the validation of the cover crop prediction model, Soil Capital randomly selected 108 fields with cover crops in the winter of 2019-2020 belonging to six farmers in the Hesbaye region. Winter wheat, potatoes, and sugar beet are the main crops at these farms. Soil Capital has a different classification for tillage practice compared to the CA & TA database (section 2.2.1 & section 2.2.2). They distinguish between conventional and conservation tillage. We compared 64 overlapping fields between the two datasets. 13 non-inversion tillage fields in the CA data set are perfectly aligned with Soil Capital's conservation tillage fields. The remaining 50 occasional inversion or inversion fields in the

CA & TA data set can also be matched and are displayed as conventional tillage in the Soil Capital database. Only one inversion field in the TA data set is shown as conservation tillage in the Soil Capital data set. Given this high match ratio, we perform external validation using Soil Capital's tillage data, that includes 114 fields with conservation tillage and 636 fields with conventional tillage.

3.2.3 Satellite data

The Google Earth Engine (GEE) platform was used to gather all remote sensing data, including Landsat-7, Landsat-8 and Sentinel-2 optical data, Sentinel-1 radar data and precipitation products as auxiliary data.

3.4.3.1 Optical remote sensing dataset

Before uploading the croplands data to the GEE platform, we first removed a 15m inwards buffer zone from the edges of the fields containing mixed pixels with a signal from roads and/or adjacent fields. For the given field dataset, a defined time period (Jan 1st 2015 to Dec 31st 2020), and cloud threshold (less than 85% cloud cover), three level-2A atmospherically corrected datasets were initially acquired, obtaining Landsat-7 surface reflectance (SR) dataset (199 scenes, Jan 1st 2015 to Dec 31st 2020), the Landsat-8 SR dataset (242 scenes, Jan 1st 2015 to Dec 31st 2020) and the Sentinel-2 SR dataset (645 scenes, Mar 28th, 2017 - Dec 31st, 2020). The multispectral instrument of Sentinel-2 acquires spectra in 13 bands with both high spatial resolution (10-60m) and high temporal resolution (5-day revisit time). GEE only provides data since 2017, after the launch of Sentinel-2B. In the next step, a more rigorous cloud removal algorithm and band correction were carried out based on the image bands i.e for Landsat 7 & 8 SR data we used a cloud mask function based on a pixel quality assurance (pixel_qa) band; for Sentinel-2 we combined a cloud mask function based on a bitmask band with cloud mask information (QA60) and a filter based on Cloud Score

Index. (see for more details Zhou et al., 2022). Then the mean values of Normalized Difference Vegetation Index (NDVI) and Normalized Burn Ratio 2 (NBR2) (see Eq. 3.1-3.2) were calculated in cloud computing for each individual field. Finally, the time series of these two indexes for the six-year time period were exported as csv. files (Ujaval, 2020).

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3.1)$$

$$NBR2 = \frac{SWIR1 - SWIR2}{SWIR1 + SWIR2} \quad (3.2)$$

Where *NIR*, *RED*, *SWIR1*, *SWIR2* represent the near infrared red band, red band, first shortwave infrared band and second shortwave infrared band respectively, and the wavelengths of these bands are shown in Table 3.1.

Table 3.1. Band wavelengths of Landsat-7, Landsat-8, and Sentinel-2 satellite sensors (W: Wavelength; R: Resolution).

Band	Landsat-7		Landsat-8		Sentinel-2	
	W (mm)	R (m)	W (mm)	R (m)	W (mm)	R (m)
Red	0.63-0.69	30	0.64-0.67	30	0.65-0.68	10
NIR	0.77-0.90	30	0.85-0.88	30	0.79-0.90	10
SWIR1	1.55-1.75	30	1.57-1.65	30	1.57-1.66	20
SWIR2	2.09-2.35	30	2.11-2.29	30	2.10-2.28	20

3.4.3.2 Sentinel-1 Radar dataset

The Sentinel-1 constellation collects data from dual-polarization C-band Synthetic Aperture Radar (SAR) sensors operating at 5.405GHz, clouds have no effect on imaging because of microwave spectrum characteristics. The GEE cloud platform provides Sentinel-1 Ground Range Detection (GRD) scenes, processing them using the Sentinel-1 Toolbox to produce calibrated, orthogonally corrected products. These GRD images were acquired in interferometric wideswath (IW) mapping mode, with a spatial resolution of 10m and a temporal resolution of 6 to 12 days.

After limiting the time period (Jan 1st 2015 to Dec 31st 2020), we obtained 4095 scenes in total of the backscatter coefficient at single Vertical-Vertical (VV) and Vertical-Horizontal (VH) polarization modes, with the sensor acquisition mode IW. We exported the time series data of VV and VH bands as csv. files. From 2015 to 2020, the average available images per field in the optical remote sensing dataset and radar dataset were 295 and 495 scenes respectively.

3.4.3.3 Precipitation data

We used the Global Satellite Mapping of Precipitation (GSMaP, Okamoto et al., 2005), which is accessible on the GEE platform. This data product was created in Japan for the Global Precipitation Measurement (GPM) mission (Skofronick-Jackson et al., 2017). It provides global hourly rainfall rates at a resolution of 0.1 x 0.1 degrees, with data available since March 2014. We compiled daily precipitation data from the hourly data on the GEE cloud platform.

3.2.4 Defining indicators for the three principles of CA

In order to more accurately quantitatively identify and characterize the three fundamental criteria of CA, seven variables were created, i.e., one to four indicators for each principle.

3.2.4.1 Crop species diversification

Based on the agricultural census (section 2.2.3), two indicators of crop diversity were calculated. **Indicator 1**: number of different annual crops in six years; **Indicator 2**: cereal abundance in rotation represents the frequency of cereal crops in the rotation, without considering the specific cereal species (Eq.3.3-3.4).

$$\text{Indicator 1} \quad \text{Number of different annual crops} = \text{Count}(\text{unique}(\text{crop code})), \quad (3.3)$$

$$\text{Indicator 2} \quad \text{Cereal abundance} = \sum_{i=1}^n \text{Cereal or not}_i, \quad (3.4)$$

Where i represents the year i of the total n years (here, $n = 6$).

3.2.4.2 Length of the cover crop growing season and periods of bare soil or fallow

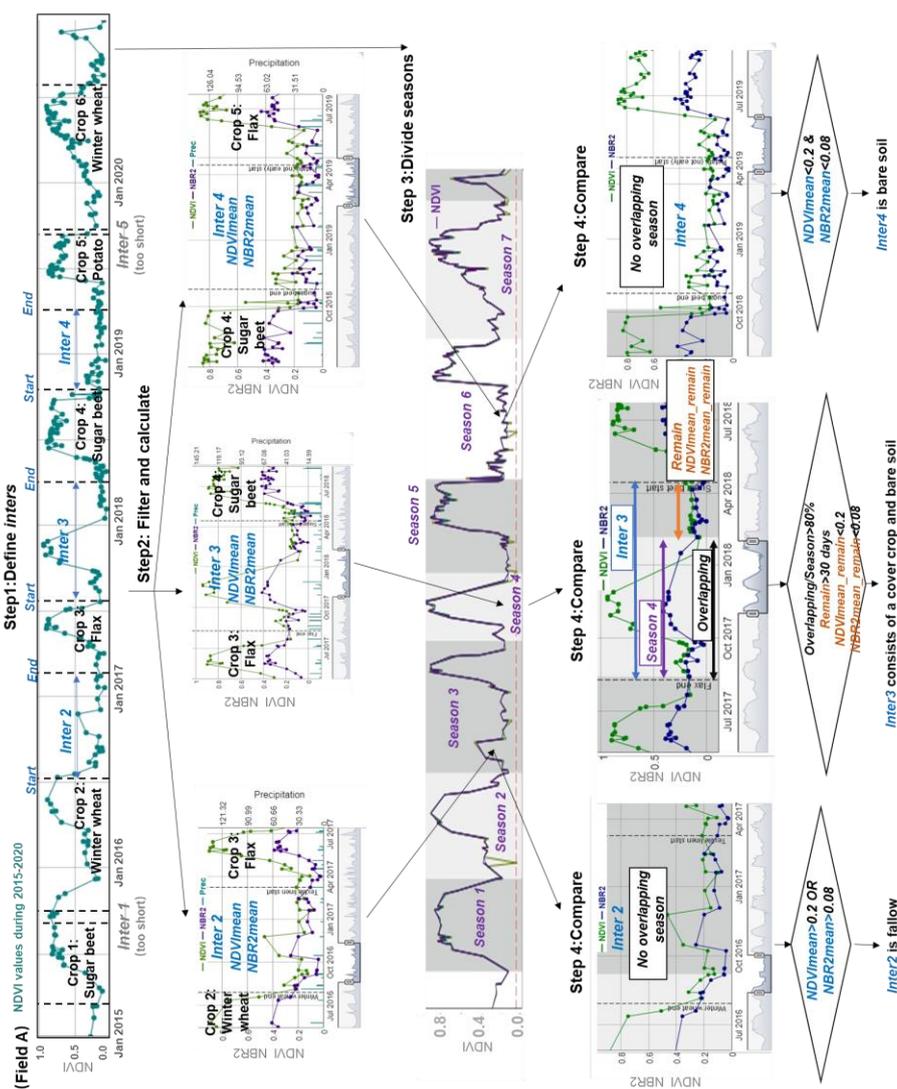
We define the period between the planting of annual main crops (as declared in the agricultural census) as *inter*, which is the time between the harvest date of the previous year's annual crop and the seeding date of the next year's crop. During this *inter* period, the field can be classified into three states: under cover crops (field that is growing cover crops), bare (field without vegetation), and fallow (field without cover crops but not bare, usually characterized by the presence of crop residues, regrowth of previous annual crops, and weeds). To determine the field status and duration, data on time series NDVI, NBR2, and precipitation, as well as annual crop types and seeding and harvest dates from 2015 to 2020, were required.

Information on the phenology of multiple annual crops (e.g. seeding, emergence, flower, maturity, harvest) is available on the website of the German National Meteorological Service (<https://www.dwd.de/>), recorded on a yearly basis by region. The Rheinland-Pfalz region of Germany, which borders Belgium at the same latitude, served as the primary reference for the seeding and harvest date. Additionally, farming and agriculture websites were used to gather information on the seeding and harvest dates for the remaining crops (Table A.1).

Since NBR2 values are significantly higher in moist soils compared to dry ones (Dvorakova et al., 2023), we introduced precipitation data into our calculations to eliminate the problem of elevated NBR2 due to precipitation. In addition, cover crops frequently do not completely overlap with *inter*. Therefore, the growing seasons obtained from automatic NDVI splitting (phenofit package, Kong et al., 2022; Kong, 2023) is utilized to assess the presence or absence of cover crops and to calculate the length of the cover crop growing season. Alternatively, an *inter* may have a single cover crop, or a combination of cover crop &

bare soil or cover crop & fallow. Based on the NDVI and NBR2 values we determined whether soil was left bare or in fallow. We then computed the total cover crop duration as well as bare soil and fallow duration over the entire observation period.

The specific steps of this procedure are listed as follows and three examples are given in Fig.3.2.



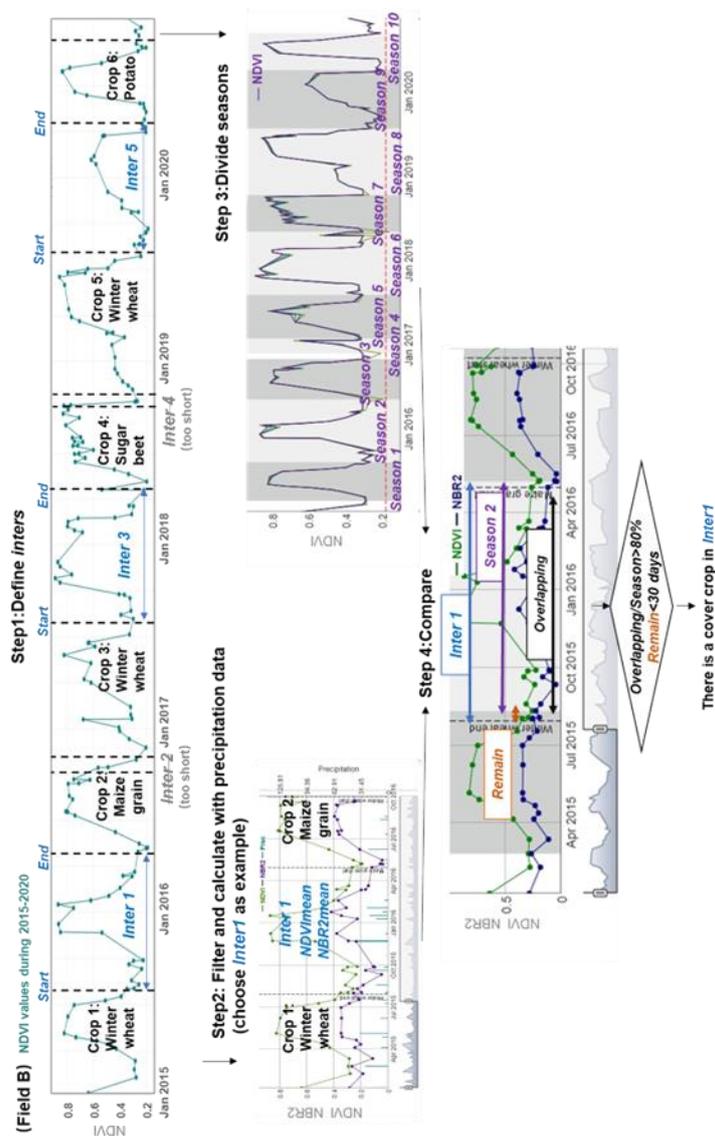


Fig. 3.2. Examples of identifying cover crops and bare soil based on long-term series NDVI and NBR2 values during 2015 to 2020. Field A: This field was fallow between the two annual crops from 2016 to 2017 (Inter2) and consisted of a cover crop and bare soil between the two annual crops from 2017 to 2018 (Inter3), while the next year (inter4) was bare soil. Field B: Cover crop was planted between the annual crops of 2015 and 2016 (inter1).

Step 1: Identify the inter period. Confirm the start (Start) and end (End) date, as well as the duration (Days) of each inter, via the harvest date (same as Start) of the previous crop and the seeding date (same as End) of the next crop. Ignore this inter if the duration is shorter than 30 days. If not, proceed to step 2.

Step 2: Preprocessing remote sensing data. The outlier points were removed using the phenofit package (Kong, 2023), followed by the calculation of the mean NDVI (NDVI_{mean}) within the inter. The NBR2 points were filtered using the precipitation data, the total rainfall over the preceding five days (Prec) of each NBR2 data point acquisition time was calculated and the data point was removed if Prec is greater than 2.5 mm, then the mean NBR2 value (NBR2_{mean}) was calculated.

Step 3: Define the growing seasons based on NDVI curves. The seasons are automatically divided by the phenofit package based on NDVI time series data (Kong et al., 2022). We ensure that it is a complete growing season of an annual crop or cover crop, rather than the spontaneous regrowth of the previous annual crop or weed, by conditioning that the NDVI peak is 0.6 or more for each season. For detailed parameter settings, please refer to the code in the data availability section.

Step 4: Determine whether a cover crop is present. Compare the inter with all seasons and compare the inter period with the growing seasons and calculate the duration of their overlap (e.g. the overlapping part in Fig. 3.2 Field A Step 4). An overlap rate (overlapping/growing season) greater than 80% indicates strong crop emergence during the inter, and we identify it as a cover crop in this study, moving on to step 6. Go to step 5 if there is no cover crop.

Step 5: Provided cover crops are absent, if NDVI_{mean}<0.2 and NBR2_{mean}<0.08, the soil remains bare during this period, set inter as

the duration of the bare soil. Otherwise, soil is fallow, set inter as the fallow period.

Step 6: If a cover crop is detected, we firstly calculate the length of the remaining period without cover crop (Remain) in the inter. If Remain is less than 30, we consider that only cover crops exist in that interval, record the number (one inter exists once), the duration, and the start and end dates of cover crop, the fallow period and bare soil is set to 0; if Remain is greater than 30, proceed to Step 7.

Step 7: Provided cover crops are present, if $NDVI_{mean} < 0.2$ and $NBR2_{mean} < 0.08$, the soil state during this period consists of cover crops and bare soil, set Remain as the duration of the bare soil. Otherwise, the soil state is under cover crop and fallow, set Remain as the fallow period.

Finally, we can obtain the number of cover crop per inter, as well as the duration of cover crop, fallow, and bare soil. In total, six years were used and the values within the five inters were summed to produce the following four indicators: **Indicator 3**: Number of the cover crops; **Indicator 4**: Duration of the cover crops; **Indicator 5**: Bare soil period; **Indicator 6**: Fallow period (Eq. 3.5-3.8).

$$\text{Indicator 3} \quad \text{Number of the cover crops} = \sum_{j=1}^k \frac{\text{cover crop presence or absence}_j}{(\text{Presence (1)} / \text{Absence (0)})}, \quad (3.5)$$

$$\text{Indicator 4} \quad \text{Duration of the cover crops} = \sum_{j=1}^k \text{Days of cover crop}_j, \quad (3.6)$$

$$\text{Indicator 5} \quad \text{Bare soil period} = \sum_{j=1}^k \text{Days of bare soil}_j, \quad (3.7)$$

$$\text{Indicator 6} \quad \text{Fallow period} = \sum_{j=1}^k \text{Days of fallow}_j, \quad (3.8)$$

Where j represents the period j of the total k inters (in this study, $k = 5$).

3.2.4.3 Tillage practices

In order to quantify the principles of mechanical soil disturbance for Hesbaye region, we utilized data on tillage practices combined with remote sensing data to develop a tillage model. The tillage model was designed to accommodate three scenarios (Short Interval, No Cover Crops, With Cover Crops) for the *inter* periods and modeled them separately. We binarized the tillage practice data of all fields (247 CA fields and 105 TA fields) and recorded them as non-inversion tillage and inversion tillage, which served as dependent variable of the model. The non-inversion tillage dataset comes from the CA fields where they implement non-inversion tillage and direct seeding; the inversion tillage dataset combines the occasional inversion tillage fields under CA and the inversion tillage fields under TA. The mean/minimum values of NBR2, NDVI, VV and VH within each *inter*, and the cover crop/bare soil duration ratio were used as independent variables of the model.

The precise steps are listed below with an example of the workflow shown in Fig.3.3:

1. Short Interval (*inter* <30 days): The interval between two annual crops is extremely short (less than 30 days). This often occurs when the next year's crop is a winter crop, and we omit this situation and categorize this tillage practice as "NA". Ploughing may occur within this interval, but the time frame between two main crops is very short, with farmers typically completing soil preparation and planting within a few days. Hence it is not possible to obtain sufficient information within such a short period from satellite imagery.
2. No Cover Crops (*inter* >30 days): In the absence of cover crops, build a model (Model I) using the mean/minimum values of NBR2, NDVI, VV and VH backscatter.
3. With Cover Crops (*inter* >30 days): In the presence of cover crops, use the proportion of cover crop/bare soil duration for the entire

interval, as well as the mean/minimum values of NBR2, NDVI, VV and VH in the remaining period to build a model (Model II).

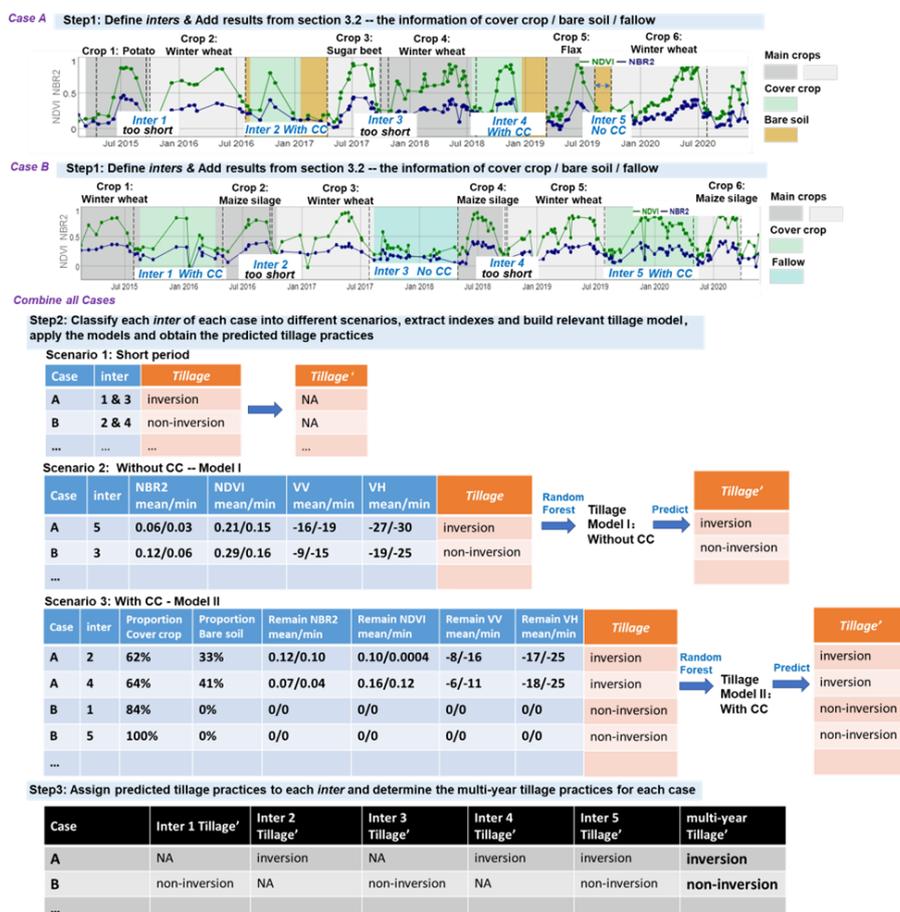


Fig. 3.3 Workflow for building tillage models and predicting non-inversion/inversion tillage practice under three scenarios (Here: in orange tables of Step 3, Tillage represents the actual tillage practices and Tillage' represents predicted tillage practices using Model I & Model II. The procedure of random forest and prediction in Step3 used 10-fold cross validation; here we simplify this part in this figure).

Each *inter* of the field is considered independently, analyzed under the corresponding scenario to obtain the predictions. Finally, we analyzed and predicted tillage practices of five different *inters* over six years for each field. The overall tillage practice of the field is defined

as **Indicator 7: Tillage practice** (Eq. 3.9). It is worth noting that the multi-year tillage practice was judged to be inversion as long as inversion tillage occurred at least once during the six-year period.

$$\text{Indicator 7} \quad \text{Tillage practice} = \text{inversion or non} - \text{inversion tillage}, \quad (3.9)$$

3.2.5 Model development and evaluation

3.2.5.1 Statistical analysis

To assess the significance of the differences between the constructed indicator values under CA and TA, we employed the Wilcoxon test. The Wilcoxon rank sum test is a non-parametric test that serves as an alternative to the independent two-sample t-test, when the data do not conform to a normal distribution. This procedure is implemented using *rstatix* and *ggsignif* R package.

3.2.5.2 Covariates decorrelation and recursive feature elimination

De-correlation analysis can be used to reduce redundant information. Only covariates with correlation coefficients less than 0.85 (Poggio et al., 2021) with all other covariates were kept, and the first covariate was removed from the set of covariates that exceeded this threshold.

Recursive feature elimination (RFE, Guyon et al., 2002) is an effective method for simplifying the structure of machine learning models, reducing dependencies and collinearity while retaining model accuracy. It is simplified by recursively eliminating features from the training dataset to select an optimal set of features. Also, RFE is frequently utilized in spatial mapping exercises with random forest models (Kuhn & Johnson 2019; Poggio et al., 2021). The algorithm first fits the model using the entire set of covariates, then evaluates its performance by cross validation while computing the importance score

of the covariates. The least important covariates are then removed from the pool, the model is fitted and evaluated again per loop. The process is repeated until the number of covariates reaches the specified number. The optimal subset of covariates after this backward selection is used to train the final model. The implementation of RFE in R can be realized with caret R package.

3.2.5.3 Classification model

Random forest is made up of numerous individual classification and regression trees that work together as an ensemble (Breiman, 2001). The fundamental principle of random forest is that each decision tree grows independently by randomly sampling replacements from the initial training dataset, with the unselected samples being used as "out-of-bag" data to assess the precision of the predictions. Without any pruning, each tree in the model will reach its maximum growth and generate a prediction, the class with the highest votes or average value of all fitted trees becomes the final prediction. The random forest model also provides importance of the model feature, which is measured by calculating the increase in prediction error (MSE) after the removal of that feature. The implementation of random forest in R can be realized with *randomForest* R package.

3.2.5.4 Accuracy evaluation

To evaluate the performance of classification, we generated a confusion matrix based on the results of 10-fold cross-validation, from which we calculated the overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and Kappa coefficient (Eq.3.10-3.13). The OA refers to the probability that the classification for each unknown point is consistent with the measured class of that point, assesses the overall performance of correctly classified. The PA illustrates how accurately the producer (or classifier) identifies objects of a specific class on the ground. While the UA describes how correct the user (or

analyst) is in interpreting the results. Kappa coefficient characterizes the agreement between the predicted and measured classification maps.

Table 3.2. Composition of the confusion matrix (NN represents the number of fields considered conventional by both measured and predicted data, CC represents the number of fields considered conservation by both measured and predicted data, CN represents the number of fields that are measured as conservation but predicted as conventional, NC represents the number of fields that are measured as conventional but predicted as conservation.)

		Reference Data		
		Conventional	Conservation	Total
Classified Data	Conventional	NN	CN	NN+CN
	Conservation	NC	CC	NC+CC
	Total	NN+NC	CN+CC	

$$\text{Overall Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{NN+CC}{NN+CC+CN+NC} \quad (3.10)$$

$$\text{Producer's accuracy} = \frac{\text{Number of correctly predictions of a class}}{\text{Total number of observations of that class}} = \frac{NN}{NN+NC} \text{ OR } \frac{CC}{CN+CC} \quad (3.11)$$

$$\text{User's accuracy} = \frac{\text{Number of correctly predictions of a class}}{\text{Total number of predictions of that class}} = \frac{NN}{NN+CN} \text{ OR } \frac{CC}{NC+CC} \quad (3.12)$$

$$\text{Kappa} = \frac{\text{Observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}} = \frac{\left(\frac{NN+CC}{NN+CC+CN+NC}\right) - \left(\frac{NN+NC}{NN+CC+CN+NC} \times \frac{NN+CN}{NN+CC+CN+NC} + \frac{NC+CC}{NN+CC+CN+NC} \times \frac{CN+CC}{NN+CC+CN+NC}\right)}{1 - \left(\frac{NN+NC}{NN+CC+CN+NC} \times \frac{NN+CN}{NN+CC+CN+NC} + \frac{NC+CC}{NN+CC+CN+NC} \times \frac{CN+CC}{NN+CC+CN+NC}\right)} \quad (3.13)$$

3.3 Results

3.3.1 Species diversification

The number of cereals and the number of different crop types are significantly different under TA and CA (Fig. 3.4). According to the standard raised by FAO (fao.org/conservation-agriculture/), CA fields

should involve at least three different annual crops. In our database, most CA fields (81%) have at least four different annual crops over six years.

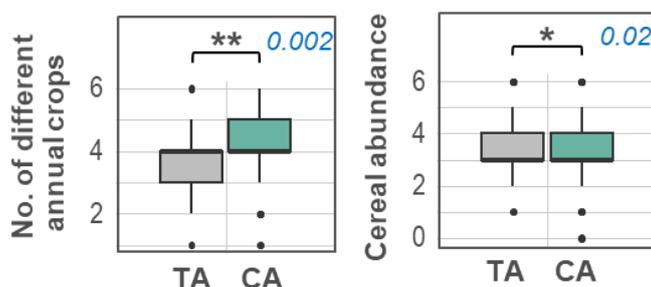


Fig. 3.4. Grouped boxplots with significance notation under Wilcoxon test, comparing the levels of two indicators related with species diversification under two farming system in Walloon region during 2015-2020. (ns: not statistically significant ($P > 0.05$), *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$); (Here: CA: Conservation Agriculture; TA: Conventional Agriculture). The blue numbers in each subplot indicate P value of Wilcoxon test.

3.3.2 Accuracy for identifying cover crops and soil cover

After comparing our classified cover crop maps with the observed data from Soil Capital, a confusion matrix was obtained with an overall accuracy of 86% (Table 3.3). For the cover crops, there is a certain probability (26%) of omission error from the producer's perspective. Soil Capital found that the omitted fields tend to have a long fallow period, which is probably caused by an inaccurate growing season classification. Nevertheless, from the user's perspective, the accuracy for the cover crops classification was 95%, which is extremely trustworthy.

Table 3.3. Confusion matrix of cover crop classification (Where CC means Cover Crop)

		Reference Data			User's accuracy
		No CC	CC	Totals	
Classified Data	No CC	56	13	69	81%
	CC	2	37	39	95%
	Totals	58	50	108	
Producer's accuracy		96%	74%		Overall: 86%

We then created box plots for the indicators related with cover crop and soil cover (Fig. 3.5) under the two farming systems and calculated the significance of the differences. Significant differences under conventional and conservation agriculture were found in the length of bare soil and cover crops.

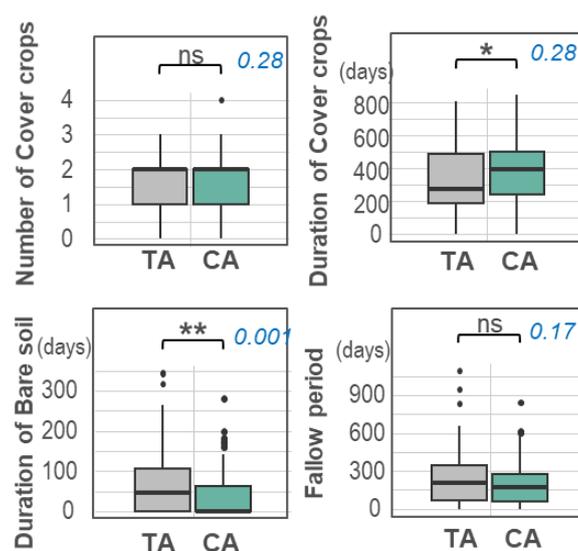


Fig. 3.5. Grouped boxplots with significance notation under Wilcoxon test, comparing the levels of 4 indicators related with soil cover under two farming system in Walloon region during 2015-2020 (ns: not statistically significant ($P > 0.05$), *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$); (Here: CA: Conservation Agriculture; TA: Conventional Agriculture). The blue numbers in each subplot indicate P value of Wilcoxon test.

3.3.3 A model of non-inversion/inversion tillage practice prediction

The main explanatory variables for the tillage model are NBR2 and backscatter coefficient, as shown by the importance of the covariates in model I and II (Fig. 3.6). In addition, the proportion of cover crop duration in the entire interval is also critical, and we found a statistically significant difference in this index between fields using non-inversion tillage and those using inversion tillage (Fig. A.2.1).

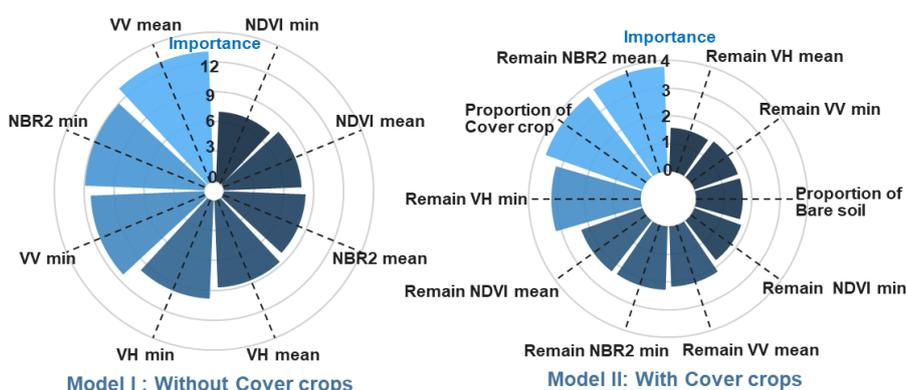


Fig. 3.6. Importance of covariates in predicting tillage practices under Model I: Without cover crops and Model II: With cover crops.

The tillage model has an overall accuracy of 66% compared to dataset from the interviews for 10-fold cross-validation and 76% compared to the external validation dataset of Soil Capital (Table 3.4). To identify inversion tillage, our user’s accuracy (87%) and producer’s accuracy (85%) were good based on external validation, but many of the non-inversion tillage fields were wrongly classified as inversion tillage.

Table 3.4. Confusion matrix of non-inversion/inversion tillage mode

10-fold cross-validation	Reference Data				User’s accuracy
		non-inversion	inversion	Totals	
Classified Data	non-inversion	112	34	146	78%
	inversion	83	114	197	58%
	Totals	195	148	343	
Producer’s accuracy		57%	78%		Overall: 66%
External validation	Reference Data				User’s accuracy
		non-inversion	inversion	Totals	
Classified Data	non-inversion	34	97	131	26%
	inversion	80	539	619	87%
	Totals	114	636	750	
Producer’s accuracy		30%	85%		Overall: 76%

3.3.4 Covariates selection and model accuracy for conservation agriculture field classification model

Ultimately, the data available for model development consisted of 247 fields in CA versus 105 fields in TA. The seven indicators (Eq. 3.3-3.9) include number of different annual crops, cereal abundance, number of cover crops, duration of the cover crop, bare soil period, fallow period and tillage practice (outputs from the tillage model). Strong correlations between two covariates were found between number of cover crops and duration of the cover crop (0.93, Fig. A.2.2). We simply need to keep one of the covariates in every pair to have sufficient information, and the number of cover crops was removed.

The recursive feature elimination results demonstrate that the best precision is obtained when four covariates are considered (Fig.3.7 (A)) and they were ranked in order of importance (Fig.3.7 (B)) as follows: Tillage practice, number of different annual crops, period of fallow and bare soil.

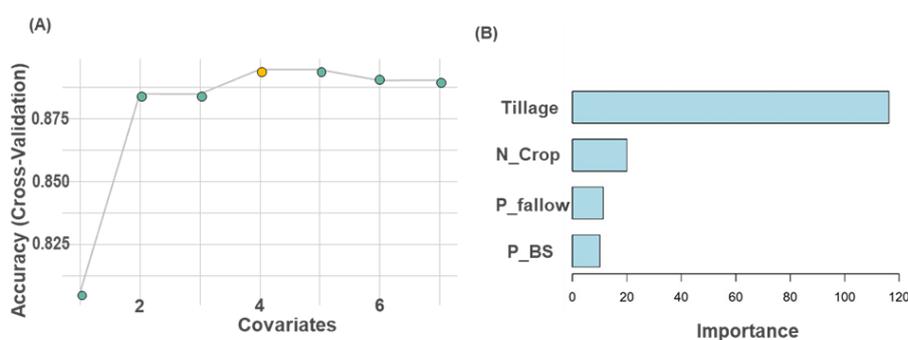


Fig. 3.7. (A) Model accuracy evolution for different number of covariates under recursive feature elimination and (B) importance of covariates. (Tillage: the tillage practice; N_crop: number of crop types; P_BS: bare soil period; P_fallow: fallow period)

The confusion matrix and model accuracy for the classification model obtained from 10-fold cross-validation are shown in Table 3.5. Random Forest, as an ensemble machine learning strategy, produces an

acceptable classification accuracy with overall accuracy of 92%. For identifying CA fields, user’s accuracy was 94% and producer’s accuracy was 95%.

Table 3.5. Confusion matrix of field classification model

		Reference Data			User’s accuracy
		Conventional	Conservation	Total	
Classified Data	Conventional	91	13	104	88%
	Conservation	14	224	238	94%
	Totals	105	237	342	
Producer’s accuracy		87%	95%	Overall: 92% Kappa:0.82	

3.3.5 Classification map for conservation agriculture

Finally, we applied the random forest model to predict and categorize 18,516 cropland fields in the Hesbaye region into conventional or conservation systems (Fig. 3.8).

The model predicts that 15.5% of these fields (2,875) were designated for CA, and 84.5% (15,641) were classified as TA. CA covers 148.2 km² and TA 830.7 km² of the Hesbaye croplands, with average field sizes of 5.1 ha for conservation and 5.3 ha for TA. Meanwhile, we zoomed in on two typical areas (Fig. 3.8). CA is practiced on a larger portion of the cropland in region B1, whereas region B2 is dominated by TA. Comparing the grading of the two most important variables for the classification model (tillage practice and the number of crop types) between the two regions, it can be noted that CA fields tend to more frequently adopt non-inversion tillage and grow more diversified crops.

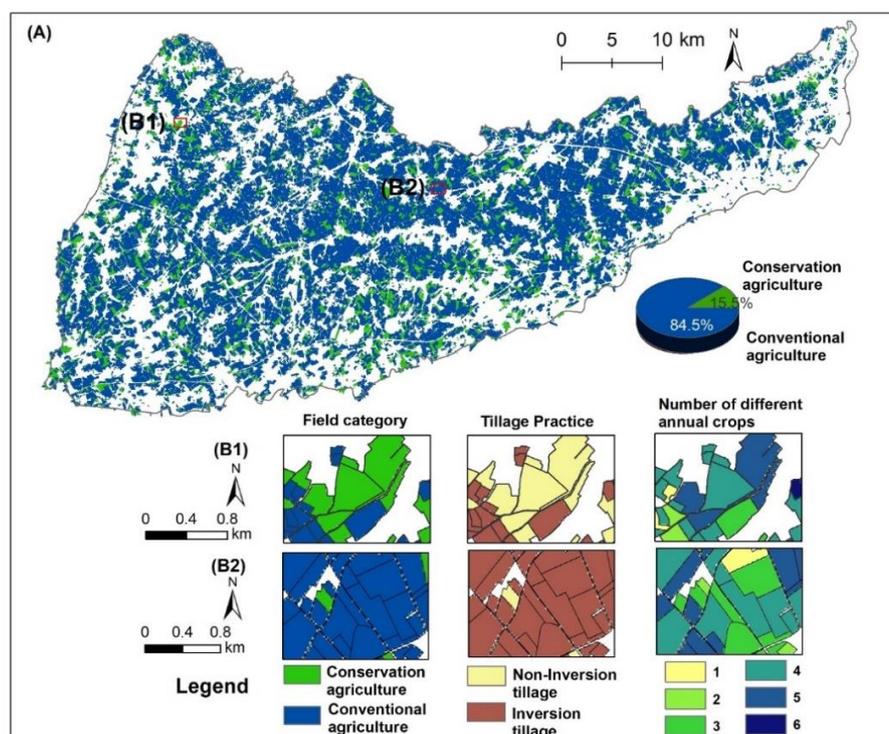


Fig. 3.8. Classified map (A) and enlarged views in representative regions are presented in (B1) and (B2), with the two most important indicators, tillage practice and the number of crop types (for the period 2015-2020), only croplands are mapped leaving grassland, forest and built up areas in white.

3.4 Discussion

3.4.1 Track management practices with time-series imagery

Time-series of optical satellite images have proven to be effective in identifying surface conditions (Laamrani et al., 2020; KC et al., 2021). The phenology of crops changes during the year, which is particularly evident in optical images. Previous studies have primarily relied on calculating the vegetation index during a fixed time period, commonly during winter months, to assess crop cover. For instance, Laamrani et al. (2020) used seasonal multi-temporal Landsat-8 satellite imagery

with linear spectral decomposition techniques to estimate crop cover in Canada during the non-growing season (November to May). Fendrich et al. (2023) used Sentinel-1 data to calculate cross-polarization ratio (CR) and generated a 100 m resolution cover crop map. Although their user's accuracy for cover crops was low (59.2%), this was the first attempt to produce a European-scale map of cover crops.

We achieved a better user's accuracy of 95% for cover crop at a finer resolution (30 m), and our automated processing workflow allows for simultaneous calculating duration of bare soil or fallow. It is worth noting that wet conditions darken soil and crop cover, making it difficult to distinguish between them (Sharma et al., 2020; Dvorakova et al., 2023). We included precipitation data to limit the effect of soil moisture when distinguishing between bare soil or fallow, but irrigation factors are difficult to characterize. In addition, we use the definition of "fallow", because so far, we can't distinguish between crop residues, regrowth of previous crops, or weedy field. To address this issue and reduce the interference of green vegetation on residue determination, future work will include implementing an NDVI threshold to restrict analysis to fields with low-vegetation conditions (e.g., $NDVI < 0.3$ as suggested by Lamb et al., 2022; Hively et al., 2018) or by analyzing trends in the NDVI and NBR2 curves over time.

Our approach matches the growing seasons divided based on NDVI data with the period without annual crops to calculate indicators for determining cover crop, which is novel compared to previous studies using fixed time periods (Thieme et al, 2020; Barnes et al., 2021). Our method enables to specifically identify the start and end date of each cover crop together with the duration of the cover crops. Moreover, our model covers five growing seasons (Fig. A.2.3), which gives a more comprehensive assessment of the implementation of cover crops than a single-year cover crop map such as produced by Fendrich et al. (2023). After all, cover crops can't be planted during the growing seasons of

winter cereals (Fig.3.3 Panel A: Inter 1&5, Panel B: Inter 2&4) and therefore a timeframe corresponding more or less to the length of the rotation will be a more representative proxy for coverage by crops and biomass input into the soil.

Differences in the spectral reflectance of crop residues can be used to characterize tillage. In addition, radar images can penetrate clouds and capture data in adverse weather conditions, making them useful for monitoring tillage practices (Kontgis et al., 2017). However, studies detecting tillage are often limited to a few fields with the same crop type for which some are ploughed (nearly) coinciding with the satellite overpass or to cropping systems with a specific period for tillage. For instance, in the tillage classification model of Sharma et al. (2020), only remote sensing data from June each year were acquired. This approach may work in very organized and structured farming contexts such as those in the United States or Latin America. But in the EU context, the smaller and more diverse fields make it difficult, and such a narrow timeframe is not suitable for detecting tillage in a rotation with summer and winter crops that are seeded in different periods. Therefore, based on the soil cover condition and crop calendar information, we targeted potential tillage periods at different dates each year. All these periods were assigned into three scenarios (short interval, with cover crops, no cover crops), analyzed and predicted separately. NBR2 and VV/VH backscatter coefficients were important variables (Fig. 3.6). This is consistent with most papers stating that the tillage index NDTI (also referred to as NBR2) is the most effective indicator for broadband satellites (Zheng et al., 2014), and that the backscatter coefficient enables detecting tillage-induced changes in soil roughness (Satalino et al., 2018).

Farmers sometimes complete ploughing, soil preparation and seeding in one day when planting a winter crop (examples in Appendix A. of Buysse et al., 2017), making it impossible to simulate tillage

practice using remote sensing data with a temporal resolution of 5-10 days. Although the accuracy of the tillage model (76%) based on external validation is not very high, we can see that the user's accuracy (i.e. the accuracy that an end-user will obtain) for inversion tillage is not bad (87%; Table 3.4). This accuracy will help monitoring the adherence to conservation farming or carbon farming schemes for which refraining from tillage is a condition. In addition, the predicted tillage map (Fig. A.2.4) show 89.9% of the fields adopted inversion tillage at least once in the past five years, and 66.8% of the tillage occurred before planting sugar beet and potato. This is because farmers prefer tillage practices before seeding root crops to improve soil aeration for the regular development of the tubers (Djaman et al., 2022). However, there are a few studies show that no-till produced similar yield and quality compared to conventional tillage (Keshavarz Afshar et al., 2019; Holmstrom & Carter, 2000).

3.4.2 Current status for classifying conservation agriculture

For their comprehensive review, Ahmed et al. (2023) selected 68 articles on CA and grouped them into five categories of conservation practices: cover crop (15), crop residues (26), crop rotation (9), mulching (4) and tillage practices (14). They noticed that these studies focused on a single facet of conservation practices at the time. We attempted to classify CA and TA fields by constructing quantitative indicators for three recognized criteria of CA, i.e. crop diversity, maximum soil cover and minimum soil disturbance. Overall, we analyzed data over a six-year period from 2015-2020. Combining data from the LPIS with time-series of optical and radar imagery and precipitation data, we obtained a classification model with an overall accuracy of 92%. To date, there is no standard for classification

accuracy for CA fields. Thomlinson et al. (1999) set an overall accuracy target of 85% for land use patterns, with no subcategories falling below 70% accuracy. Our model successfully achieved this target.

Contrasting with our predicted CA adoption rate of 15.5 %, The European CA Federation (ECAAF; <https://ecaf.org/>) calculates a significantly lower conservation agriculture adoption rate for Belgium at only 0.03%. There may be an overestimation of our CA fields, as our initial data collection lacked field audits. Nevertheless, the indicators established in our study are clearly different under two farming system, with CA fields overall scoring better on the indicators (Fig. 3.8).

Moreover, there is an ambitious aspiration to develop regenerative agriculture (Tittonell et al., 2022) focusing on restoration and rebuilding of the soil health and planet's ecosystems and helping to meet the objectives of the EU Soil Strategy for 2030 (Rehberger et al., 2023; European Commission, 2021). There are five principles of regenerative agriculture policy, in addition to the three principles that overlap with CA, the other two principles that were not discussed here include minimizing the use of chemical inputs and integrating livestock. There are currently some studies that use hyperspectral drones or aerial photographs to identify livestock species and count animals on a small scale (Ocholla et al., 2024). However, using satellite data to detect livestock or chemical inputs on a larger scale remains challenging. A feasible solution is to convert administrative-level statistical data into pixel-scale distribution using spatial models or based on crop distribution. (Liu et al., 2024; Shen et al., 2022).

3.4.3 Potential applications and limitations of our work

Some data products which can be used as input data for soil carbon models are generated simultaneously (downloadable in Data availability). For instance, the RothC model (Coleman & Jenkinson,

2014) utilizes the degree of soil cover (bare or vegetated). Integrating these data into soil carbon models enables credible and robust simulations of soil organic carbon dynamics, and supports the development of Monitoring, Reporting, and Verification (MRV) platform (COWI, 2021) to ultimately validate the effectiveness of soil carbon sequestration as a strategy for greenhouse gas removal.

One major limitation of this paper is this study relies on high quality and high-frequency remote sensing data covering the length of a crop rotation (3-5 years). Bad weather conditions such as cloud cover will cause data loss, affecting the final forecast. To alleviate this problem, future research may interpolate time series dataset to fill the gaps (Yan and Roy, 2020). Real time detection of management practices is still not possible, as the method relies on time-series.

3.5 Conclusion

By combining long-term series of optical and radar satellite data with precipitation data, and using machine learning methods, a conservation agricultural fields classification model in the Walloon region of Belgium was successfully established (overall accuracy of 92%). The input indicators of this model cover the three principles of CA. The model can distinguish fields under conservation agricultural land and associated management practices for large areas. The adoption rate of conservation agriculture in the Hesbaye region is 15.5%, and the dominant management practice that determines the predicted results is the tillage practice. However, tillage followed by seeding in one or two days and management practices such as application of fertilizers and other chemical input remain beyond our current ability to model through remote sensing.

Data availability: We provided the source codes and five cover crop maps during 2015-2020 at field scale:
<https://doi.org/10.5281/zenodo.10370700>

Chapter 4

Simulating SOC Changes at Lonzée Station Using the RothC Model

4.0 Outline

Agricultural soils have great potential for carbon sequestration. Although soil organic carbon (SOC) dynamics models such as RothC model have frequently used to simulate the changes in SOC resulting in CO₂ exchanges with the atmosphere, the effects of replacing measured boundary conditions with readily available data from e.g. remote sensing has received less attention. This is nonetheless an important step for running the model runs to cover fields for which detailed measurements do not exist or upscaling model runs to cover an entire region. Based on ground truth data from a long-term experimental site in Lonzée with regular crop rotations, we ran the RothC model on Belgian cropland to simulate SOC stock changes from 2007 to 2017. All the boundary conditions required by the model were replaced one by one with remote sensing or agricultural data to evaluate models' uncertainty under different assimilated inputs. In addition, a cover crop identification and biomass prediction procedure were developed for calculating carbon input from cover crops, which are often ignored in current RothC model simulations. Ultimately, models using either all in-situ data (bias of -0.16 t·ha⁻¹) or all remote sensing data (bias of 1.76 t C·ha⁻¹) both achieved good accuracy. The Lonzée station maintained the same SOC stock level after 10 years

cultivation due to the significant fertilization effect. Planting cover crops instead of leaving the field fallow during the winter also had a positive effect on SOC stock. The largest model uncertainty from replacing measured data with remote sensing comes from the carbon input calculation for main crops and cover crops, while climate and soil data obtained through remote sensing introduced almost no errors. Our work is a key step in assessing the uncertainties when applying the model to larger regional, national, and even global scales for which detailed boundary conditions are not available.

4.1 Introduction

Soils constitute the largest terrestrial carbon reservoir (~1500 Pg C), and even minor changes in this reservoir can have a significant impact on atmospheric CO₂ levels and the global carbon balance (Batjes, 1996; Smith, 2008). However, the soil organic carbon (SOC) stocks of European croplands are experiencing severe declines (Ferreira et al., 2022; Panagos et al., 2018). The EU Soil Strategy emphasizes targeted and sustainable soil management practices as crucial for increasing soil carbon in agricultural lands by 2030 (European Commission, 2021). Enhancing soil carbon sequestration in agriculture not only contributes to climate change mitigation, but also supports achieving climate neutrality (Chabbi et al., 2017; Lal, 2004ab).

Process-based soil models can estimate trends and changes in SOC over time and assess how different management practices affect carbon sequestration (De Rosa et al., 2023). The RothC model (Coleman & Jenkinson, 1996), also known as the Rothamsted Carbon model, is one of the most commonly used first-order kinetic SOC model for simulating carbon turnover in non-waterlogged topsoil. The crucial advantages of the RothC model include its requirement for a small amount of input data and its simple and transparent model

structure compared to other carbon models (Lesschen et al., 2020). The model has been validated under different ecosystems, climate and land use types (Geremew et al., 2024; Maas & Lal, 2023; Fantin et al., 2022; Senapati et al., 2014; Yokozawa et al., 2009; Kaonga & Coleman; 2008; Shirato & Taniyama; 2003).

However, most of these studies are confined to the field scale where long-term experiments are available, since the RothC model requires specific management information such as carbon inputs from crop residues and farmyard manure. The availability of long-term experiments worldwide is limited. If we could acquire the necessary boundary conditions for the RothC model from remote sensing or other sources that are readily available, leveraging its advantages of large-scale coverage, cost-effectiveness, and temporal consistency (Weiss et al., 2020; Alvarez-Vanhard et al., 2021), the model could either be run for any field. Eventually, the model runs for all fields could provide a means of monitoring SOC dynamics at the landscape or regional scale as proposed in the monitoring, reporting and verification (MRV) framework of Smith et al. (2019).

Until now, several studies have applied the RothC model at regional or global scales (Zhang et al., 2024; Jordon et al., 2022; FAO, 2020; Morais et al., 2019). Additionally, RothC has been integrated into the JULES land surface model and is used as part of the JULES Earth system models (Clark et al., 2011). These studies provide mature options for most boundary conditions, utilizing remote sensing imagery and national or global scale products. For example, historical monthly climate products such as the Climatic Research Unit (CRU) dataset at a resolution of 0.5° produced by the University of East Anglia Climatic Research Unit (2020); the ERA5 Monthly Aggregates dataset at 0.5° resolution (Bell et al., 2021); and TerraClimate (~ 4 km; Abatzoglou et al., 2018). We selected TerraClimate due to its high spatial resolution. Clay content and initial SOC stock should ideally be

obtained from national datasets when available, as they are based on local conditions. If not, global soil products, such as the Global Soil Organic Carbon Map (GSOC; FAO & ITPS, 2018), the Harmonized World Soil Database (HWSD; FAO & IISA, 2023), or SoilGrids (Poggio et al., 2021) can be utilized. These maps are constructed based on legacy soil profile data, including WOSIS (Batjes et al., 2017), LUCAS (Tóth & Montanarella, 2013), and AfSIS (Leenaars, 2013), with soil samples collected over a long period: 1900-2016. To align with the baseline SOC sampling year, we employ digital soil mapping methods (Chen et al., 2022) to precisely predict baseline SOC contents. For the soil cover indicator, FAO (2020) proposed using 16-day NDVI products (from the MOD13A26 product, 1 km) and set the threshold at 0.6, which means that soil is considered covered above this value. We can also adopt NDVI threshold methods but use higher resolution satellites (Landsat series and Sentinel-2) and choose our own threshold value based on in-situ conditions.

However, the estimation of carbon inputs in current regional studies are still coarse, which is a key primary limitation for these simulations (Morais et al., 2019). For instance, starting from a final SOC stock, Jordon et al., (2022) and Zhang et al., (2024) ran RothC in inverse mode to estimate carbon inputs, while FAO (2020) and Smith et al., (2005) used estimated Net Primary Productivity (NPP) as carbon inputs. In Belgium, we benefit from an open-access database on annual crops and field layouts i.e. the land parcel information system (LPIS) provided by the Walloon Region (<https://geoportail.wallonie.be/>). This allows us to use allometric functions (Keel et al., 2017) and crop yield information to obtain more accurate carbon inputs similar to the ones used in long term experiments.

Furthermore, an increasing number of fields are adopting conservation agriculture practices, including the effective

implementation of cover crops during winter months (Ahmed et al., 2024). After all, maintaining a minimum soil cover during winter in particular for erosion prone areas is one of the conditionalities for payment to farmers in the Common Agricultural Policy (2023-2027; <https://agriculture.wallonie.be/accueil>). Unfortunately, the current agricultural census only includes main crops and does not account for cover crops, thus underestimating the relatively high potential of cover crop for carbon sequestration (Poeplau & Don, 2015; Mazzoncini et al., 2011). Recent advancements have utilized remote sensing data to predict cover crop presence (Fendrich et al., 2023; Barnes et al., 2021) and estimate aboveground biomass (which determines carbon input) using vegetation indices or Leaf Area Index (LAI) (Kharel et al., 2023; Goffart et al., 2021; Dong et al., 2020). Building on this foundation, we can integrate these methodologies to develop scalable and replicable workflows for estimating carbon inputs from cover crops.

Therefore, the overall goal of this study is to provide all boundary conditions using remote sensing data or statistical data for RothC models. These remote sensing-derived boundary conditions are a key step that can expand field-scale models to the landscape level. The specific work that needs to be done can be summarized as follows:

1. Validate the RothC model in Belgian cropland using in-situ data.
2. Identify alternative boundary conditions for the RothC model from remote sensing or agricultural data, in particular develop a procedure to assess the presence of cover crops in crop rotations and predict the aboveground biomass of cover crops using satellite data.
3. Evaluate the uncertainties and model performance gaps introduced by replacing in-situ data with remote sensing data.

4.2 Materials and Methods

4.2.1 Site description

The Lonzée Terrestrial Observatory, located at 50°33'05.7"N, 4°44'46.1"E in Belgium, is recognized for having one of the longest and most complete data series on cropland in Europe. It has served as a node in both the CarboEurope-IP and FLUXNET networks since 2004, collecting flux and meteorological data (Pastorello et al., 2020). Lonzée station officially integrated into the ICOS (Integrated Carbon Observation System) network in November 2017.

The observatory is located on a relatively flat plateau and covers an area of approximately 12 ha (Buysse et al., 2017). The climate of the site is temperate maritime, with a long-term average annual temperature of 10.8°C and an average annual precipitation of 680 mm.

This station basically operates a four-year crop rotation system, alternating between sugar beet, winter wheat, potato, and winter wheat. Soil is basically ploughed to a depth of about 30 cm. The soil is classified as Luvisol according to the FAO classification.

4.2.2 RothC Model

The RothC model (Coleman & Jenkinson, 1996) is a first-order kinetic model that simulates carbon turnover in topsoil, incorporating the effects of temperature, moisture, and plant cover. It divides SOC into five pools: four active pools, which include decomposable plant material (DPM), resistant plant material (RPM), microbial biomass (BIO), and humified organic matter (HUM), and one inert pool, inert organic matter (IOM). For each active C compartment, At the beginning of the month, the C pool size is Y_0 , and it will decompose to Y_t at the end of the month, following the exponential decay function (Eq.4.1):

$$Y_t = Y_0 \times (1 - e^{-abckt}) \quad (4.1)$$

Where k is the decomposition rate constant for each pool (default values from Coleman et al., 2024: DPM: 10; RPM: 0.3; BIO: 0.66; HUM:0.02 in years⁻¹), t is 1 / 12, since k is based on a yearly decomposition rate. Additionally, the decomposition rate is affected by climate and soil cover conditions: a , b , and c are the rate modifying factors for temperature, moisture, and soil cover, respectively.

The decomposed C material ($Y_0 - Y_t$) of each compartment within the month will split between (BIO+HUM) and CO₂, with the proportion decided by clay content. In addition, incoming external carbon source will split between DPM and RPM (for plant residue) or between DPM, RPM and HUM (for manure), based on set ratio parameters.

4.2.3 Boundary Conditions

In this section, we prepared the necessary boundary conditions (Table A.3.1) for the RothC model, detailing both in-situ and remote sensing-derived data sources for each input.

4.2.3.1 Climate data

We obtained monthly precipitation and temperature data for the Lonzée site from open-access ICOS data portal (<https://data.icos-cp.eu/portal/>). Potential evapotranspiration is calculated using the Thornthwaite equation (Thornthwaite, 1948), based on the average monthly temperature, the average length of day (in hours) and the number of days in a month. Then potential evapotranspiration is converted to open-pan evaporation by dividing by 0.75.

As an alternative remote sensing input data, we choose TerraClimate (Abatzoglou et al., 2018) because it is easily accessible (from the public catalog of Google Earth Engine), has long-term coverage (from 1958 to now), and relatively high spatial resolution (~4 km). TerraClimate provides monthly temperature and rainfall datasets directly, where we calculate the average of the minimum and maximum temperatures as the mean temperature.

4.2.3.2 Soil data

Loncée station conducted soil sampling twice. In August 2007 and in September to October 2017, soil samples were collected and soil properties were estimated at 5 standard depth intervals (0-5cm, 5-15cm, 15-30cm, 30-60cm, 60-100cm)

([https://traitementsinfosol.pages.mia.inra.fr/icos/BE-](https://traitementsinfosol.pages.mia.inra.fr/icos/BE-LonCarbonReportv2.html)

[LonCarbonReportv2.html](https://traitementsinfosol.pages.mia.inra.fr/icos/BE-LonCarbonReportv2.html)). SOC stock calculated following Eq. 4.2:

$$SOC\ stock = SOC\ content \times BD \times depth \quad (4.2)$$

the SOC stock in the 0-30 cm soil layer, was measured as $46.7 \pm 1.9\ t\cdot ha^{-1}$ (Buysse et al., 2017), and in September to October 2017, measured as $42.5 \pm 2.6\ t\cdot ha^{-1}$ (SABY et al., 2024). To obtain baseline and final SOC stock values for model running and validation, we converted SOC stocks from the two sampling campaigns in 2007 and 2017 using the equivalent soil mass method. (Wiesmeier et al., 2015) The converted SOC stocks were $40.00 \pm 1.6\ t\cdot ha^{-1}$ and $39.98 \pm 2.5\ t\cdot ha^{-1}$, respectively.

The initial SOC stock value in 2007 was additionally predicted using the Digital Soil Mapping (DSM) method. Using the random forest model developed for the prediction of SOC contents in the Walloon Region (see Zhou et al., 2022 for more details on the model covariates and development), the predicted SOC content in 2007 was obtained by replacing model's covariates with the corresponding values for 2007. The values of bulk density (BD) can be calculated using various available pedotransfer functions (Hollis et al., 2011; Xu et al., 2015). For this study, we chose to use the one developed by Manrique & Jones (1991) based on SOC content (%) (Eq.4.3).

$$BD = 1.660 - 0.318 \times SOC\ content^{1/2} \quad (4.3)$$

The model output indicates the same mean value as the measured SOC stock, at $46.7 \pm 4.2\ t\cdot ha^{-1}$, but with a larger standard deviation range. After conversion using the same ratio, the equivalent mass SOC stock would be $40.00 \pm 3.4\ t\cdot ha^{-1}$.

For the clay content, we used the value (11.2%) from soil sampling measurements in 2017. We assumed that clay contents remained constant over the study time period (2002-2017). For the alternative option, we obtained publicly available clay products from the website of the Walloon Region (<https://geoportail.wallonie.be/catalogue/e90eb7cf-8f7d-40ab-9df9-5c34ddf387ea.html>), with a clay content of 13.6% at the Lonzée station.

4.2.3.3 Soil cover

The RothC model needs to identify whether the soil is bare or vegetated as soil decomposition rates are higher when the soil is bare. However, it does not provide a comprehensive definition of the level or degree of “covered.” For the Lonzée station, estimated time-series above ground biomass (AGB) data, exported biomass during harvest (yield), and post-harvest residual AGB datasets are available. Soil was identified as covered when the field had an AGB record; otherwise, the field was considered bare.

We utilized the Normalized Difference Vegetation Index (NDVI), which widely recognized for assessing vegetation greenness and density, to distinguish between bare soil and vegetation cover. However, up to now, the unified NDVI threshold has not been established for the RothC model. Here, we established a specific threshold for the Lonzée site. Long time-series NDVI data from Landsat 5/7/8 and Sentinel-2 were downloaded from Google Earth Engine (GEE). The binary monthly soil cover indicator was derived by iterating NDVI threshold values from 0.2 to 0.65 in increments of 0.05. This indicator was compared with the soil bare/covered binary indicator determined by the AGB values, generating a confusion matrix. Overall accuracy analysis identified an optimal NDVI threshold of 0.35 (Fig. A.3.1).

4.2.3.4 Carbon input from main crops

The Bolinder Equation (Bolinder et al., 2007) is a widely used allometric equation to estimate the soil C input by relating the yield to the amount of above- and belowground plant material. It describes the total amount of the plant C returned to soil (C_i) as consisting of four components: main harvested crop products (C_p); straw, stover and other post-harvest aboveground residues (C_s); root tissue (C_R); and from extra-root material (C_E). C_p (Eq.4.4) equals the dry matter yield multiplied by the C content of plant (CC , usually 0.45 is used), and the C in the other three fractions (C_s, C_R, C_E) can be calculated based on C_p using crop-specific allocation ratio values R (Eq.4.5-4.7). The carbon from these four fractions, multiplied by the proportion of each fraction that is returned to soil, gives the total input C_i . (Eq.4.8):

$$C_p = \text{yield} (DM) \times CC \quad (4.4)$$

$$C_s = \frac{R_s}{R_p} \times C_p \quad (4.5)$$

$$C_R = \frac{R_R}{R_p} \times C_p \quad (4.6)$$

$$C_E = \frac{R_E}{R_p} \times C_p \quad (4.7)$$

$$C_i = (C_p \times S_p) + (C_s \times S_s) + (C_R \times S_R) + (C_E \times S_E) \quad (4.8)$$

Where R_p, R_s, R_R, R_E are the relative plant C allocation coefficients and S_p, S_s, S_R, S_E are the corresponding proportions of each fraction returned to the soil. (Table A.3.5)

Bolinder et al. (2007) used a reference depth of 40 cm for belowground residues. We need to scale the belowground inputs to a soil depth of 30 cm to align with our study (Eq. A.3.1).

The Lonzée station has documented the annual yield information, which allowed us to calculate C_p then get total C_i use the Bolinder

equation. When we have detailed by-products and post-harvest AGB values, we can also set it as C_s and calculate C input (C_i) based on this.

Alternatively, if measured data are not available, annual crop information and field layouts can be collected from the Walloon region website (<https://geoportail.wallonie.be/>). The annual yield information of different crop types for each province can be found from the Belgian Statistical Office (<https://statbel.fgov.be/>). It is important to note that these productions are given as fresh matter weight, and we need to convert them to dry matter yield by multiplying the dry matter content proportion parameters (parameters can be found in Hendriks et al., 2023). The annual C input obtained based on statistical data generally aligns with the C input obtained based on measured value ($R^2 = 0.86$, Fig. A.3.2).

To convert the annual C input into the monthly C input data required by the model, 50% of the residue is allocated to the harvest month, and the remaining 50% is allocated equally to the three months before harvest (Morais et al., 2019).

4.2.3.5 Carbon input from cover crops

Since 2009, cover crops (mustard, oats) are common for the Lonzée trial and their crop cycle starts after winter wheat. The species of cover crop, sowing date, termination date, and time-series of AGB values have been recorded. Cover crops are not harvested but are often used as a green manure, with the entire biomass being incorporated into the soil during tillage (Poeplau & Don, 2015; Schaefer et al., 2020). Thus, the entire biomass (AGB and roots) is used as a source of C input to the soil. The root and extra-root material are calculated based on the ratio parameters (R_R and R_E , Table A.3.5). The total biomass multiplied by the C content of the plant material (using 0.45 as before) gives the C input from cover crops, which is added as a single pulse in the month of cover crop termination.

Calculated C input from cover crops using remote sensing data involves two steps:

Identify Cover Crops: The main idea is to compare the crop growing seasons documented in the census database with the growing seasons derived from remote sensing data (using the *phenofit* package in R; Kong, 2023), to determine if there is a full growing season between main crops. The detailed mechanism can be found in Zhou et al. (2024), and specific identification procedure for the Lonzée station is described in Appendix C. Following this procedure, we can identify whether cover crops are present or absent, as well as determine their sowing and termination dates.

Estimate AGB of Cover Crops: We initially calculated the LAI using the Enhanced Vegetation Index (EVI), obtained from GEE, utilizing data from Landsat 5/7/8 and Sentinel-2 (Eq. 4.9 & 4.10, Boegh et al., 2002):

$$EVI = 2.5 * ((NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1)) \quad (4.9)$$

$$LAI = 3.618 \times EVI - 0.118 \quad (4.10)$$

Then, we estimate the AGB of cover crops by establishing a regression model between time-series AGB data and the LAI. The linear model equation expressed as (Eq. 4.11), with a R^2 of 0.73 and an RMSE of 0.62 (Fig. A.3.3).

$$AGB = 0.27 \times LAI - 0.19 \quad (4.11)$$

After predicting AGB for all available RS data points, we calculated the maximum AGB in each interval between the cover crop's NDVI peak date and the end date obtained during the cover crop identification process, using it as the final AGB of the cover crop that goes to the soil. This is because NDVI does not completely correspond to biomass, especially during the senescence stage when NDVI decreases but AGB remains unchanged (Burger et al., 2024).

We calculated belowground biomass and carbon input using the same process as above.

4.2.3.6 Farmyard manure

The Lonzée site added farmyard manure twice during our ten-year study interval. 28 t·ha⁻¹ fresh matter was added in August 2009, containing 3.04 t C·ha⁻¹; and 50 t·ha⁻¹ fresh matter in August 2015, containing 5.43 t C·ha⁻¹.

Manure application information for individual fields is difficult to obtain, we can estimate it with the total amount of manure for the whole region (city/province/country) and the regional agricultural land area. First, annual manure amount (fresh matter) of each livestock type (t·year⁻¹) is calculated (Eq.4.12). Then, we need to convert the wet weight to C content (Eq.4.13) based on C proportion under different manure management systems (such as solid storage, liquid slurry, on-pasture) for different livestock.

Annual manure per livestock type

$$= \text{Number of livestock per type (head)} \times \text{Manure per livestock (t} \cdot \text{head}^{-1} \cdot \text{year}^{-1}) \quad (4.12)$$

Total C content of manure (4.13)

$$= \sum_i^n \left(\sum_j^m (\text{Manure}_i \times \text{Prop}_{ij} \times C_{ij}) \right)$$

Where i represents each type of livestock (total n types); j represents each management system (total m systems). Manure_i is the annual manure production of livestock type i (From Eq.4.12); Prop_{ij} is the proportion of the manure (%) under management system j for livestock type i ; C_{ij} is the carbon content ratio (t·t⁻¹) of the wet weight manure under management system j for livestock type i .

Finally, combine the utilized agricultural area (UAA) to get the farmyard manure C input (Eq.14).

$$\begin{aligned} \text{Input C from farmyard manure (t C} \cdot \text{ha}^{-1}) & \quad (4.14) \\ & = \frac{\text{Total C content of manure (t C)}}{\text{UAA (ha)}} \end{aligned}$$

The number of each livestock type and UAA for each province is available from the Belgian Statistical office (<https://statbel.fgov.be/>). Estimated manure for different livestock types can be summarized from the literature (Table A.3.2, Scarlat et al., 2018). In Volume 4, 'Agriculture, Forestry and Other Land Use,' of the 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (IPCC, 2019; Tables 10A.6-10A.9), the proportion of various manure management systems for each type of livestock is provided (Table A.3.3). Parameters for the C content ratio of manure (wet weight) can also be found in previous work (Table A.3.4, Hendriks et al., 2023).

4.2.4 Workflow

We first initialized the RothC model by starting with an empty SOC stock and ran it for 1000 years until it reached equilibrium, matching the SOC stock measured in 2007. The spin up process used multi-year monthly average climate data provided by the Lonzée station from 1989 to 2006. Since we use the default model configurations for all parameters, which were calibrated with long-term field experiments, we directly validated the RothC model on Belgian cropland by using the monthly in-situ boundary conditions from 2007 to 2017 (Table.4.1: Model 0) and comparing the final model output with the measured SOC stock in 2017.

Then, we substituted the in-situ boundary conditions one by one with remote sensing or agricultural data (Table.4.1: Model 1-7), and compared the model accuracy and dynamic changes in SOC stock under different input data combinations. The final simulation run with all remote sensing-derived boundary conditions (Table.4.1: Model 8).

The entire modelling process is implemented in R. All of the decomposition modification rates were calculated using functions in

the package SoilR (Sierra et al., 2014). For the main running part of the RothC model and calculation of each input data, we built new functions to better calculate data and modify the model. These codes are available online [<https://doi.org/10.5281/zenodo.13625934>].

Table 4.1. Summary of the simulations under different input data combinations (Green cell represents in-situ data; Blue cell represents remote sensing or agricultural data.)

Model 0	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 1	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 2	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 3	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 4	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 5	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 6	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 7	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM
Model 8	Climate	Clay	Base period SOC	Soil cover	C input main crop	C input Cover crop	FYM

4.3 Results and Discussions

4.3.1 Changes in SOC stock

We simulated monthly SOC stock changes for Lonzée station during 2007-2017 (Fig. 4.1). It changes a lot during the ten years, varies from 35 to 45 t·ha⁻¹. The final predicted SOC stock is 39.82 t·ha⁻¹, which corresponds well with the measured SOC value, 39.98 ± 2.5 t·ha⁻¹. There are two obvious peak values caused by the addition of FYM, totaling 8.47 t C·ha⁻¹. The neutral situation, where the final and baseline SOC are almost similar.

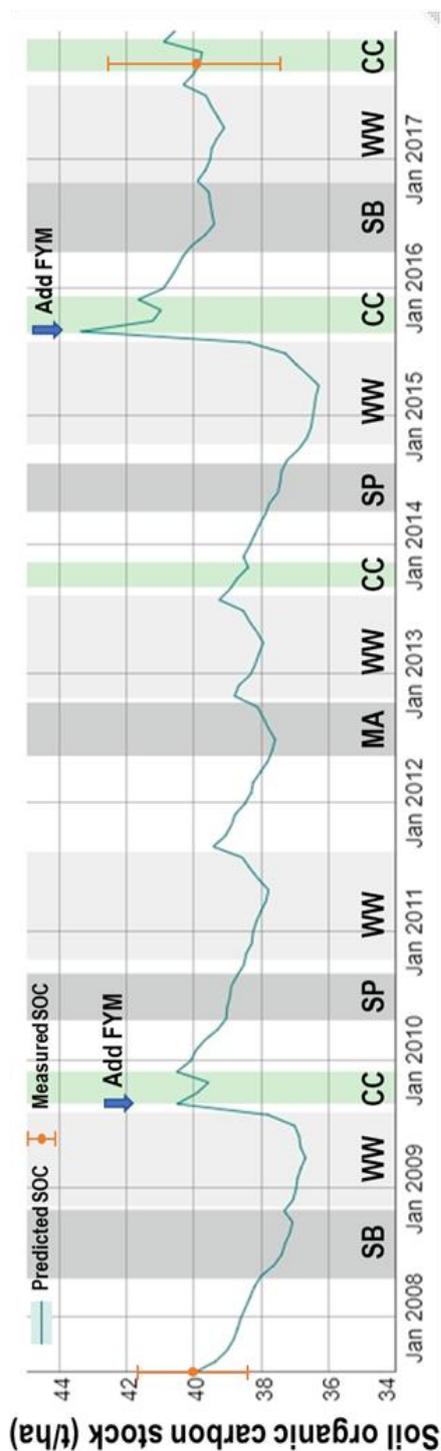


Fig. 4.1. Predicted SOC stock by RothC model. Orange points and their error bars represent the mean and SD values of SOCS observed in the two sampling years. The grey background indicates the main crops during 2007-2017 and the green background indicates the cover crop between main crops. Here, SB: sugar beet, WW: winter wheat, SP: seed potato, MA: maize, CC: cover crop.

Please note that the measured SOC stock in 2017 is till rather high as a result of FYM applied after the harvest of winter wheat in autumn 2016. Otherwise the global trend shows an overall decrease. Bolinder et al., (2020) also found similar result that manure applications have the highest impact on increasing SOC compared to other management practices based on twenty reviews using paired comparisons. This is because manure contain carbon (C)-rich substrate and improves net primary production, which brings more biomass and crop residues (Huang et al., 2022; Hijbeek et al., 2019).

Additionally, SOC stock shows increases during the years with winter cereals and silage maize (e.g., in 2009, 2011, 2013, 2015 and 2017) but decreases during the years with sugar beet and potatoes (e.g., in 2008, 2010, 2014 and 2016). Bolinder et al. (2015) pointed out that below-ground residue inputs to soil for root crops is much lower compared to other annual crops. The INTERREG Carbon Farming project (2021) selected the top five most promising measures to protect the climate through carbon farming, mentioning that crop rotations should be enriched to improve SOC stocks while reducing the area of potatoes, carrots, and other root crops in the rotation.

Leaving soil bare between two main crops (e.g., winters of 2007 and 2011) will cause a continuous decrease in SOC, but planting cover crops (e.g., in 2009, 2013, and 2015) is effective in slightly slowing these decreases. Replacing the bare fallow period in winter with cover crops can increase the SOC stock and serve as an effective measure to offset CO₂ emissions (Lal, 2004ab). Poeplau & Don (2015) estimated a potential global carbon sequestration via cover crops of 0.12 Pg C yr⁻¹, compensating for 8% of the direct annual greenhouse gas emissions from agriculture.

We validated the model using SOC stock measurements taken twice at the long-term experiment (LTE), as shown in Fig. 4.1. However, we don't have a reference to compare whether we correctly

simulated the trend of changes during this interval. Therefore, soil sampling should be repeated more frequently to provide a stronger basis for model verification. However, due to the slow rate of soil carbon change, detecting changes over a short timeframe is difficult (Raffeld et al., 2024). This is because the changes can be smaller than the measurement errors from the laboratory (e.g., precision ranges from 1.2 to 15.8% for loss-on-ignition, 1.6–4.2% for Walkley & Black, and 1.3–7.1% for dry combustion) and spatial heterogeneity (Goidts et al., 2009). According to most MRV protocols, soil re-sampling at 5-year intervals is appropriate (Raffeld et al., 2024).

A novel method to verify the C model is calculating SOC stock changes from CO₂ exchange measured in a flux tower (Smith et al., 2019). The main principle is to calculate the net ecosystem carbon budget by considering net ecosystem exchange, C inputs (e.g., fertilization), and C outputs (e.g., harvest, fires) (Smith et al., 2010; Kutsch et al., 2010). The strength of using flux data is the growing availability of flux tower data worldwide (e.g., over 1000 active and historic flux measurement sites of FLUXNET, fluxnet.org) compared to the limited number of LTE data (e.g., over 250 studies recorded in long-term soil experiments (LTSEs)). Also, the flux sites have hourly data records, allowing for model verification with short-term experiments. Buysse et al. (2017) using flux data calculated the carbon budget also for Lonzée station, and found a high C loss of 9.9 t·ha⁻¹ over 12 years (2004 to 2016), accounting for about 22% of the C stock in the topsoil. This result differs significantly from the inventory values in our study, which showed a loss of 0.02 t·ha⁻¹ during 2007-2017. The discrepancies between flux tower measurements and SOC changes may be attributed to several factors, including data acquisition and processing, necessary assumptions, and the mismatch in scale (Smith et al., 2019), also difference in microbial carbon use efficiency, which can significantly introduce uncertainty. Flux towers

are highly appropriate for detailed gas exchange measurements due to their high temporal resolution. However, for long-term changes in SOC, these measurements must be aggregated, which can introduce potential errors. Additionally, the import and export of residues, harvests, and seeds are often mentioned as factors contributing to uncertainty (Feigenwinter et al., 2023; Osborne et al., 2010).

At a long-term experimental site located just 2 km from the Lonzée site, known as the Longs Tours site (also called Liroux, in Gembloux), the effects of different management practices (residue export, farmyard manure, residue return (RR)) on SOC were studied from 1959 to 1997 (Buysse et al., 2013). Over the fifty-year period, the average SOC stock under the RR treatment showed almost no change, with a small increase of $0.07 \text{ t} \cdot \text{ha}^{-1} \cdot \text{yr}^{-1}$, which aligns more closely with our results rather than those from the flux tower. Moreover, after the first 20 years of the experiment, the SOC stocks under all three treatments became relatively equilibrium. Obviously, it is challenging to compare the results of a continuous measurement of CO_2 fluxes with a SOC stock change between two points in time.

4.3.2 Uncertainties from different inputs into the model

We substituted the in-situ boundary conditions one by one with remote sensing or agricultural census datasets. The final SOC stocks and their standard deviations (SD) under different simulations are shown in Table 4.2. The effects of data replacement on the model outputs are illustrated in Fig. 4.2, which calculated based on the difference in the final SOC stock predicted by the model.

First, replacing climate data only introduced minor uncertainties, slightly underestimating the final SOC. Comparing the coherence of climate data from site-measured values with those from climate data products (Fig. A.3.4), current temperature products fit perfectly. However, precipitation products obviously overestimate (RMSE = 29 mm) and struggle to capture extreme precipitation events. The

overestimated precipitation consequently leads to lower final SOC due to higher decomposition rates in wet soil conditions (Coleman et al., 2024). One effective approach for enhancing the accuracy of precipitation data is the application of rainfall merging algorithms (Hu et al., 2019), such as correcting biases between satellite precipitation and ground-based meteorological measurements.

Table 4.2. Predicted final SOC stock and model errors under different data input combinations. ($t\ C\cdot ha^{-1}$)

Model	Model Info	Predicted Final SOC stock	Bias (Predicted-Final)
Model 0	All measured	39.82	-0.16
Model 1	Replace climate	39.64	-0.34
Model 2	Replace clay	40.07	0.09
Model 3	Replace baseline SOC	39.82	-0.16
Model 4	Replace soil cover	39.45	-0.53
Model 5	Replace C input from main crop	41.74	1.76
Model 6	Replace CC input from Cover crop	41.01	1.03
Model 7	Replace FYM	39.35	-0.63
Model 8	All RS	41.74	1.76

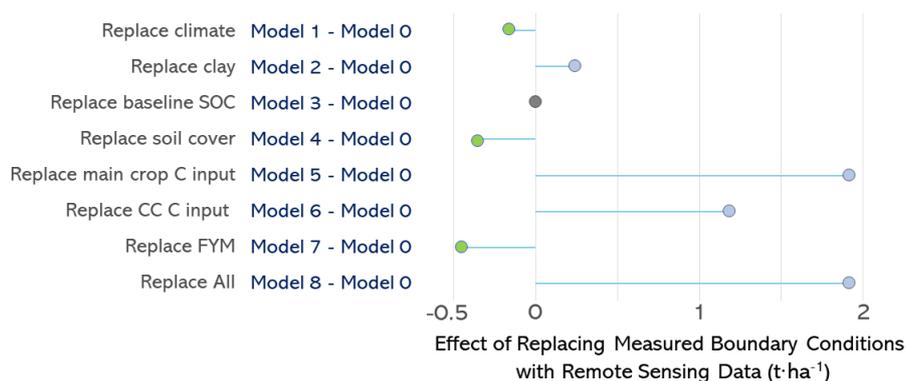


Fig. 4.1. Effect of replacing measured boundary conditions with remote sensing derived one, which calculated based on the difference under two models.

The models also show insensitivity to soil information. In current works (Jordon et al., 2022; FAO, 2020; Morais et al., 2019), baseline SOC and clay contents are often derived from available worldwide soil products, such as SoilGrids (Poggio et al., 2021), HWSD (FAO & IISA, 2023). In our study, the clay content obtained from Walloon soil products closely matched the values from soil campaigns, differing by only 2.4%. In addition, the baseline SOC did not introduce any errors to the predicted mean SOC stock value, the only difference was a higher SD ($2.66 \text{ t C}\cdot\text{ha}^{-1}$), indicating the successful construction of our robust DSM model for the Walloon region. Given these good simulation results, we recommend utilizing more region-specific soil products or developing SOC DSM models, which allow for the simulation of SOC for specific years, unlike global soil products that only provide maps for certain years.

Compared with the first three models (Model 1 – Model 3), using RS-derived soil cover info and FYM from census data have moderate negative effect ($<-0.5 \text{ t C}\cdot\text{ha}^{-1}$) on model output. We underestimated the input of FYM, since it is a complex topic, the FYM have so much different types and chemical elements, also different management methods (Doblas-Rodrigo et al., 2023). However, we did not find more detailed databases or novel methods to solve this issue. In most regional simulations, this issue is simply ignored (Jordon et al., 2022; FAO, 2020). For the soil cover indicator, most differences occur in the month of harvest when NDVI decreases. When we simply calculate the mean NDVI over a time period, bias might exist due to the uneven distribution of temporal data points.

The largest uncertainties arise from overestimating C inputs from both main crops and cover crops. First, relying on regional statistical values rather than yields measured in the field can introduce uncertainty, and residue management practices cannot be assessed using remote sensing. Additionally, different allometric equations

chosen for C inputs from main crops can introduce considerable uncertainties (Keel et al., 2017). Taghizadeh-Toosi (2016) noted that allometric functions assuming a fixed ratio between crop yield and root biomass tends to overestimate root C inputs, especially with technologies that improve crop yield. Cagnarini et al. (2019) also recommend using fixed root biomass and rhizodeposition values rather than shoot: root ratio values.

The main reason for the overestimation of C input from CC is the misidentification of CC occurrence in 2007, as shown in Fig. 4.3. After winter wheat was harvested in August 2007, the fields were not ploughed until January 2008. Winter wheat may have regrown during this period, causing an increasing NDVI trend, which was mistakenly identified as a cover crop. Additionally, the duration of cover crops in 2009 was overestimated due to insufficient NDVI points throughout the winter (Appendix C). With the availability of higher temporal resolution Sentinel-2 data since 2015, CC identification become very accurate. This indicates that despite the incorrect estimations caused by limited data in our case, the process of identifying CC can be effectively extended to more sites and for recent periods.

In addition, information on CC duration is very valuable and should be incorporated into AGB model. Studies have shown that longer CC cover duration delivered more C to the soil (Seitz et al., 2023; Hendriks et al., 2023). However, mustard grown as CC at Lonzée is frost sensitive (Gabbrielli et al., 2022) and was terminated in November and December, limiting the availability of comparative data to model AGB under different CC durations. Additionally, distinguishing CC species through remote sensing is challenging (Wang et al., 2023), which could greatly impact AGB and C input (Selin Noren et al., 2021), thus also introducing uncertainty when substituting measured data with remote sensing data.

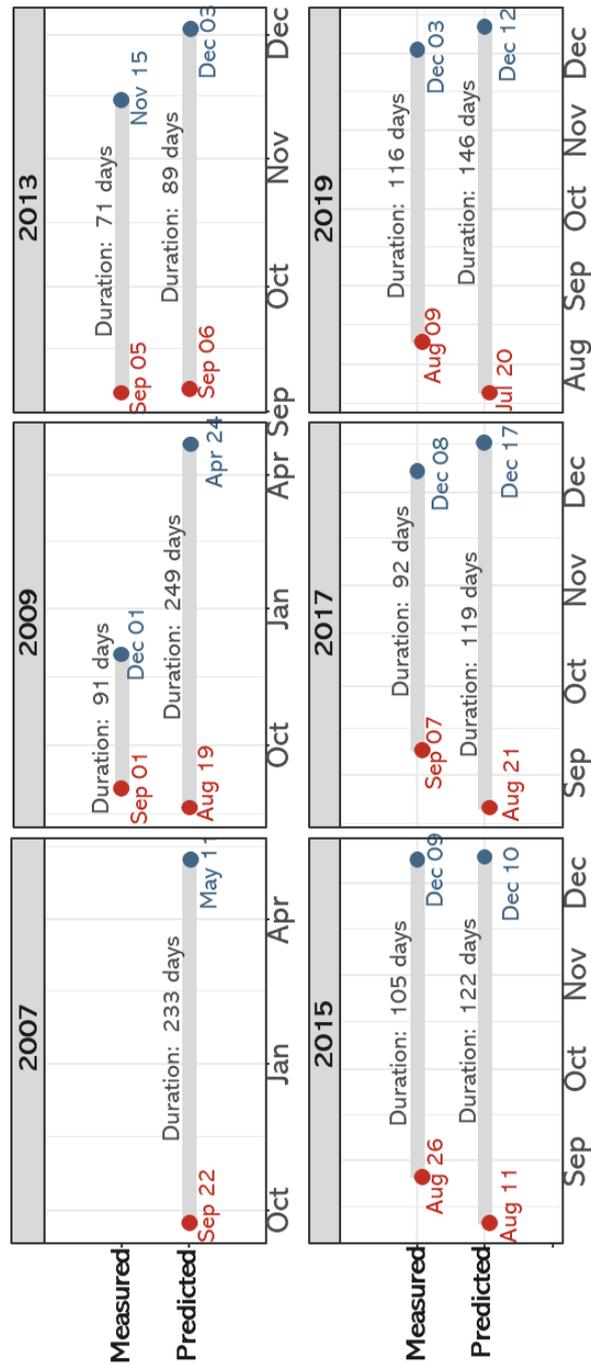


Fig. 4.2. Comparison of recorded cover crop calendar and predicted durations from satellite

The model run with all RS boundary conditions showed a bias of $1.76 \text{ t C}\cdot\text{ha}^{-1}$, corresponding to a 4.4% relative error in the final SOC stock. as some positive effects were compensated by negative ones. However, the effect of boundary conditions on SOC turnover processes have interaction and are non-linear, meaning that their effects cannot simply be summed or subtracted. But overall, remote sensing data provide a feasible alternative for carbon modeling, enabling the model running from point scale to landscape level in further steps.

4.3.3 Limitations of the work

First, when calculating the moisture rate modifying factor, only the water input from precipitation is considered, while irrigation is not included. Morais et al. (2019) estimate the irrigation amount by calculating the water requirements of specific crops, which could be a feasible solution.

Second, the soil cover rate modifying factor in the model is binary, causing drastic changes, and it is difficult to define the threshold. A rate factor that changes gradually according to the soil cover (such as based on NDVI or LAI) should be developed.

Last, the consideration of management practices is not comprehensive enough, for example, tillage practices affect SOC stock. Varvel and Wilhelm (2010) compared SOC changes under six tillage systems in LTE and found that no-till system stores more SOC than other tillage treatments in the 0-30 cm depth. In subsequent work, a rate modification factor d for tillage can be created. This factor can be obtained by parametrizing the model under LTEs that control tillage practices. For example, van Wesemael et al. (2019) compared SOC changes under three tillage treatments: conventional tillage with a moldboard plough, deep de-compaction with a heavy tine cultivator, and reduced tillage (RT) with a spring tine cultivator at a long-term

trial, which at Gentinnes, Belgium and started from 2008. In the 0-20 cm topsoil layer, reduced tillage resulted in a significant increase in C content, particularly in the $C > 20 \mu\text{m}$ fraction. Based on the graphical results, the carbon sequestration rate for RT was estimated to be approximately $0.6 \text{ t}\cdot\text{ha}^{-1}$. This value can then be used in combination with other modification rates to estimate the rate modification factor for tillage. It can also be derived from existing literature, for example, Jordon & Smith, (2022) suggests a tillage rate modifier of 0.93 for reduced-tillage, based on a systematic review database.

4.3.4 Extensions and future work

For future work, in addition to the above, we can expand our model verified at the site scale to the entire Wallon region or even Belgium. Since RothC is a soil model (Fig. 1.4), which focused on field-scale simulations, expanding to larger spatial scales requires further calculations. Our approach is to use remote sensing data to obtain boundary conditions over larger areas, including climate and management data, following the Monitoring, Reporting and Verification (MRV) framework proposed by Smith et al. (2020; Fig. 1.6). Alternatively, we could utilize global models such as ORCHIDEE, the Community Land Model (CLM), Integrated Biosphere Simulator (IBIS). These land surface models are designed to operate at broader spatial scales, but they typically run at coarser spatial resolutions (e.g. default resolution of ORCHIDEE is $0.5^\circ \times 0.5^\circ$), which may result in the loss of detailed field-level information. Balancing the need for large-scale simulations with the preservation of local details will be a critical consideration in future research.

We can also run the model under different land use patterns to simulate carbon turnover in Belgian forest or grassland systems, based on ICOS station at Vielsalm, Brasschaat, Maasmechelen and Dorinne

(<https://www.icos-cp.eu/observations/national-networks/belgium>). Additionally, evaluating the greenhouse gas trend under different management scenarios is crucial for assessing the impact of regenerative agriculture (Jordon et al., 2022) and also support MRV protocol for soil carbon.

4.4 Conclusions

We successfully validated the RothC model, which showed a bias of $-0.16 \text{ t} \cdot \text{ha}^{-1}$ for a ten-year period using data from a long-term experimental station in Lonzée, Belgium. Substituting in-situ boundary conditions with remote sensing data yielded satisfactory results ($1.76 \text{ t C} \cdot \text{ha}^{-1}$ or a relative error of 4.4% compared to the SOC stock). The substitution of measured data by remote sensing products is a first step in upscaling the model runs to the regional scale.

The primary uncertainties introduced by remote sensing data arise from the estimation of carbon inputs from both main crops and cover crops. However, the accuracy of cover crop identification improved substantially with the addition of satellite data of higher temporal and spatial resolution. This highlights the need for advancements in remote sensing products concerning crop phenology and productivity.

Chapter 5

Evaluating SOC Changes and Sequestration Potential at landscape scale

5.0 Outline

Accurate assessment of soil organic carbon (SOC) dynamics is crucial for evaluating the effectiveness of various land management practices for carbon sequestration in agricultural soils. We used the RothC model to simulate changes in SOC stocks from 2015 to 2024 across 10,102 fields in the Walloon region of Belgium. To achieve this, we prepare all required boundary conditions at the landscape scale, including climate and management data. A major advancement of this research was the use of remote sensing technology to accurately identify field management practices. In particular, the distinction between organic carbon inputs from main crops and cover crops proved to enhance the reliability of simulations on SOC dynamics. We examined the impact of input data resolution on SOC predictions by testing different spatial stratifications: per field, soil association groups, and pedoclimate zones. The results revealed that differences in input resolution could lead to entirely opposite trends. For instance, the regional SOC stock calculated based on per field is in decrease, with an average loss of $-0.38 \text{ t} \cdot \text{ha}^{-1}$ over ten years. However, it is shown as a carbon sink under the other two spatial resolutions. Furthermore, we defined three management scenarios: reduced tillage

(RT), long-term cover crops (CC) on all fields with potential for planting CC, and a combination of RT and CC (RT+CC). We then calculated their carbon sequestration potential compared to the business-as-usual (BAU) scenario, which only considered conventional agricultural fields. The results, based on field-level resolution, showed that the RT, CC, and RT+CC scenarios all effectively offset SOC losses. Specifically, the carbon sequestration rate under the CC+RT scenario was $0.22 \text{ t} \cdot \text{ha}^{-1} \cdot \text{yr}^{-1}$, increasing SOC stocks on about half of the fields. This highlights the critical importance of effectively implementing agricultural carbon sequestration practices.

5.1 Introduction

The implementation of carbon sequestration in agricultural soils is drawing increased interest because of its significant potential to mitigate climate change by removing greenhouse gases from the atmosphere (Amelung et al., 2020; Rumpel et al., 2023). However, the intensification of agricultural production has led to unsustainable soil degradation (Derpsch et al., 2024; Kopittke et al., 2019). Sanderman et al. (2017) estimated that 12,000 years of land use have resulted in a global carbon loss of 133 Pg in agricultural soils, with the rate of loss accelerating dramatically over the past 200 years. Currently, long-term experiments have demonstrated the feasibility of sustainable carbon sequestration practices for agricultural soils, such as reduced tillage, crop residue management, and cover cropping, underscoring the need for further widespread adoption (Bai et al., 2019; Page et al., 2020; Rodríguez et al., 2022).

Soil organic carbon (SOC) dynamics models are crucial tools for assessing long-term changes in SOC. These models are increasingly used to simulate the impact of agricultural management practices on

carbon sequestration in agricultural soils (Lugato et al., 2013; Smith et al., 2005). Among these models, the RothC model (Coleman and Jenkinson; 1996) is one of the most widely used for predicting carbon sequestration potential across various regions and globally (FAO, 2020; Gutierrez et al., 2023; Jordon et al., 2022; Pesce et al., 2024). Accurately measuring the carbon sequestration potential of different management practices involves two critical steps: (1) accurately assessing the baseline carbon stock, and (2) precisely defining and quantifying the management scenarios. Although SOC dynamic models have been validated for long term experiments (Romanenkov et al., 2019; Li et al., 2019), extrapolating to a region with a strong variation in soils, crops and management types still remains challenging (Smith et al. 2022).

Numerous studies have employed stratification or raster data to model the baseline carbon stock across the entire region (Wiesmeier et al., 2016; van Wesemael et al., 2010; Morais et al., 2019; Lugato et al., 2013). Stratification methods typically divide the study area into homogeneous land units (HLUs) based on factors such as climate region, soil type, soil texture, land use, or management practices, with average values per unit serving as model inputs. For example, Morais et al. (2019) defined approximately 17,000 Unique Homogeneous Territorial Units to estimate attainable SOC stocks globally. However, the potential uncertainty introduced by this aggregation process raises questions about the reliability of the overall carbon storage calculations based on aggregated inputs.

Currently, an increasing number of studies use regional raster data to directly calculate model inputs (Jordon et al., 2022; Gutierrez et al., 2023; FAO, 2020; Zhang et al., 2024). However, these studies often simply calculate C input to the soil by running RothC in inverse mode (Jordon et al., 2022; Gutierrez et al., 2023) or estimating net primary productivity (NPP) using the Miami model (FAO, 2020; Pooplau &

Dechow, 2023). The main drawback of these approaches is that they do not differentiate whether the overall carbon input originates from the main crop, cover crop, or organic amendments. This lack of specificity makes it difficult to accurately identify which management practices could be improved when defining management scenarios.

As a result, proposed management scenarios often adopt a simple proportional increase or decrease in the original C input (FAO, 2020; Jordon et al., 2022; Wiesmeier et al., 2016), without clearly explaining how this increase in carbon input can be achieved. For example, the Global Soil Organic Carbon Sequestration Potential map (GSOCseq; FAO, 2020) defines three Sustainable Soil Management practices corresponding to low, medium, and high C inputs, with respective increases of 5%, 10%, and 20%. However, the specific management methods used to achieve these increases are not detailed. Similarly, Wiesmeier et al. (2016) defined management scenarios with C inputs varying by -20% to +20% for SOC simulations in Germany without specifying the practices behind these changes. Although Jordon et al. (2022) introduced scenarios distinguishing reduced tillage, cover cropping, and ley-arable management measures, the C input values were estimated based on literature review, assuming changes in the proportion of plant residue input rather than reflecting the extent to which these practices will be adopted by farmers.

With the rapid development of remote sensing technologies, we can now accurately determine the management practices employed on each field, especially whether cover crops and inversion tillage practice are used (Chapter 3). This allows for precise field-scale predictions across landscapes, enabling us to identify and modify specific management practices that can be improved. Such accurate predictions can better inform agricultural policies for instance in the framework of the Common Agricultural Policy (CAP), by evaluating

management practices on a per-field or per-farmer basis, thus offering a more reliable basis for policy decisions and implementation.

To address these issues, the overall research goal is to accurately calculate carbon input, simulate changes in SOC stock under different spatial stratifications at the landscape level, and compare the carbon sequestration potential of different management practices. The specific objectives are:

1. Use remote sensing techniques to predict the boundary conditions required by the RothC model, with a focus on accurately calculating carbon inputs from main crops and cover crops.
2. Predict and compare baseline SOC stocks under different spatial stratification, at field scale, soil associations and pedoclimatic zones.
3. Establish scenarios based on conservation agriculture practices and compare the carbon sequestration potential of different management measures.

5.2 Methods and Materials

5.2.1 Study Area

Our study area (Fig. 5.1) is located in the northern part of the Walloon region of Belgium. Our study area fully covers eight 0.125° grid cells that correspond to those of earth surface models such as ORCHIDEE (<https://orchidee.ipsl.fr/>), including 24,277 agricultural fields. We screened the fields and retained only those that have been used as cropland every year since 2015, excluding temporary and permanent grasslands, as we do not have management data for grasslands. We also ensured that the fields area is larger than 1 ha, as smaller plots are easily affected by surrounding roads or other features that are likely to result in a mixed pixel when obtaining remote

sensing (RS) data. Finally, 10,102 agricultural fields were retained, with a total area of 50,656 ha. Field lay-out and individual field boundaries are collected from the website of the Walloon Region (<https://geoportail.wallonie.be/>), the Land Parcel Information System (LPIS). These fields are distributed across four provinces in Wallonia: Brabant Wallon, Namur, Hainaut, and Liège. The yield and livestock statistics presented in Section 5.2.3 are based on data collected at the provincial level.

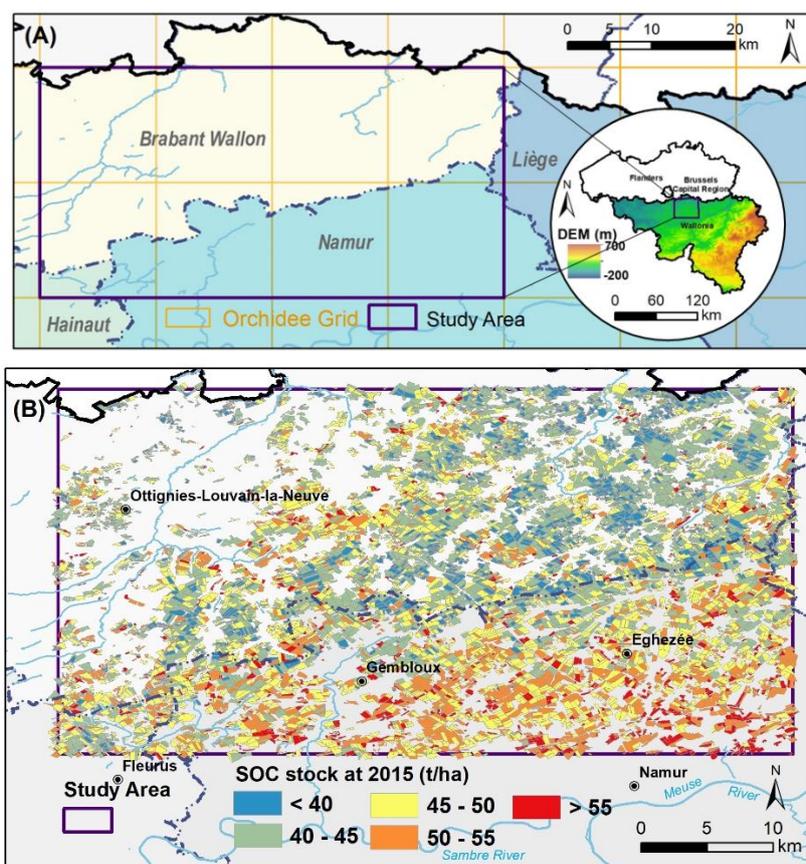


Fig. 5.1. (A) Location and extent of the study area, where orange grids represent 0.125° ORCHIDEE grids. (B) Distribution of fields within the study area and their initial SOC stock values (from section 5.2.3.1 Soil data).

5.2.2 RothC model

The RothC model, developed by Coleman and Jenkinson (1996), is a first-order kinetic model that simulates carbon turnover in the topsoil. The SOC has four active pools include Decomposable Plant Material (DPM), Resistant Plant Material (RPM), Microbial Biomass (BIO), and Humified Organic Matter (HUM) and one inert pool named Inert Organic Matter (IOM).

Carbon decomposition in each active pool follows an exponential decay function, considering the effects of temperature, moisture, and plant cover. In our study, we have extended this model by incorporating the effect of ploughing to more comprehensively account for management practices based on Jordon & Smith (2022). The decomposition process can be described by Eq. 5.1:

$$C_t = C_0 \times (1 - e^{-abcdkt}) \quad (5.1)$$

Where:

- C_0 is the initial carbon pool size at the start of the month;
- C_t is the carbon pool size at the end of the month;
- k is the decomposition rate constant for each pool (year^{-1}). Default values from Coleman et al. (2024) are: DPM: 10, RPM: 0.3, BIO: 0.66, HUM: 0.02.
- t is 1/12, as k is based on a yearly decomposition rate.
- a, b, c are rate-modifying factors for temperature, moisture, and soil cover, respectively. Detailed calculations are provided in Coleman et al. (2024).
- d is the newly introduced rate-modifying factor for ploughing, $d = 1$ for conventional tillage and $d = 0.93$ for reduced tillage

The decomposed carbon material ($C_t - C_0$) from each compartment is distributed between (BIO+HUM) and CO_2 , with the proportion determined by clay content. The carbon input from plant residues was allocated between DPM and RPM in a default ratio of 1.44:1; for manure inputs, 49% was allocated to DPM, 49% to RPM, and 2% to HUM (Coleman et al., 2024).

5.2.3 Preparing the boundary conditions

We ran the RothC model for a ten-year period, from January 2015 to December 2024. Executing the model over this timeframe requires preparing all the necessary boundary conditions (Table 5.1).

Table 5.1. Boundary conditions for RothC model and their source

	Variable	Units	Source	Spatial Resolution
Soil data	Initial carbon pool	t ₁ ha ⁻¹	Zhou et al., 2022	30 m
	Clay content of the soil	%	Geoportal of Wallonia	~100 m
	Depth of soil layer sampled	cm	Set as 30	
Climate data	Monthly mean air temperature	°C	Terra Climate	~4 km
	Monthly precipitation	mm	Terra Climate	
	Monthly open pan evaporation	mm	Calculated from monthly mean air temperature	
Soil cover	Soil cover (0 or 1)	none	Time series NDVI dataset	30 m
Management	Monthly input plant residues	t C ha ⁻¹	Calculated based on yields and Time series LAI dataset	Per field
	Monthly input farmyard manure	t C ha ⁻¹	Calculated based on livestock numbers	Per province

5.2.3.1 Soil data

The spatial distribution of initial SOC stock in the study area is shown in Fig. 5.1(B). We calculated the 30cm SOC stock using the following formula (Eq.5.2):

$$\begin{aligned}
 \text{SOC stock (t/ha)} & \\
 &= \text{SOC content}(\%) \times \text{BD}(\text{g/cm}^3) \\
 &\times \text{depth}(\text{cm})
 \end{aligned}
 \tag{5.2}$$

The SOC content was derived from the SOC map presented in Zhou et al. (2022). The sampling points used in this study were

collected between 2015 and 2019. We assumed these values to represent the model's initial conditions as of January 2015, as our primary focus is on SOC changes and sequestration potential rather than absolute values. The values of bulk density (BD) were calculated using the pedotransfer function based on SOC content (Eq. 5.3; Manrique & Jones, 1991):

$$BD = 1.660 - 0.318 \times SOC \text{ content}(\%)^{1/2} \quad (5.3)$$

The clay content data for the Walloon Region are based on soil analysis for the soil map carried out in the 1960's. The clay content map can be accessed through the Geoportal of Wallonia (<https://geoportail.wallonie.be/catalogue/e90eb7cf-8f7d-40ab-9df9-5c34ddf387ea.html>).

5.2.3.2 Climate data and soil cover

Climate data and soil cover information are used in both the spin-up and simulation steps of the model.

Regional climate data were obtained from monthly TerraClimate datasets (Abatzoglou et al., 2018), which are available on Google Earth Engine (GEE) with a spatial resolution of 4 km. While TerraClimate directly provides monthly precipitation data, monthly average temperature needs to be calculated as the mean of minimum and maximum temperature data. Open-pan evaporation (OPE) was estimated using the Thornthwaite equation (Thornthwaite, 1948) based on average monthly temperature.

For the model spin up, we used the average climate data from 1965 to 2014, representing the 50 years preceding our simulation period. During the simulation period from January 2015 to December 2023, we used actual monthly data. For the period beyond December 2023, where actual data were not yet available, we used monthly averages from previous years.

To determine soil cover conditions, we obtained long-term NDVI (Normalized Difference Vegetation Index) remote sensing data and calculated monthly means. We set a threshold of $NDVI > 0.35$ to indicate soil cover, with a corresponding c factor of 0.6. When $NDVI \leq 0.35$, soil was considered bare, with $c = 1$. This threshold of 0.35 was determined by Zhou et al. (2024) through comparison of NDVI threshold and measured above-ground biomass (AGB) at a long-term experiment station.

NDVI data were downloaded from GEE. For the spin-up period, we used monthly mean NDVI from Landsat-5 data for 1984 to 2012. For the simulation period from January 2015 to December 2018, we utilized a combination of Landsat 7, Landsat 8, and Sentinel-2 data. We limited our soil cover information to this four-year period to align with management data (section 5.2.3.3). In cases where cloud cover resulted in missing data for certain months, we interpolated values using the average of the previous and following months.

5.2.3.3 Field management: Crop rotation, cover crop, Tillage.

In our previous research (Chapter 3), we processed, analyzed, and predicted field management measures for part of the study area from 2015 to 2019. This analysis included determining whether each field employed conventional or conservation agriculture systems, whether tillage practice was generally inversion or non-inversion, as well as annual crop rotations, cover crop presence, and cover crop duration. The main results of these predictions are illustrated in Fig. 5.2. Tillage information can be directly used as a reference for rate modifier d in our model, with all non-inversion tillage considered as reduced tillage and $d = 0.93$.

Here we focused on management measures for the four-year period from 2015 to 2018. This time frame was chosen because, while there is no unified rotation pattern for all fields, four-year rotations are

more common. It's important to note that the occurrence and duration of cover crops at the end of 2018 were determined by combining main crop information and remote sensing data from 2019.

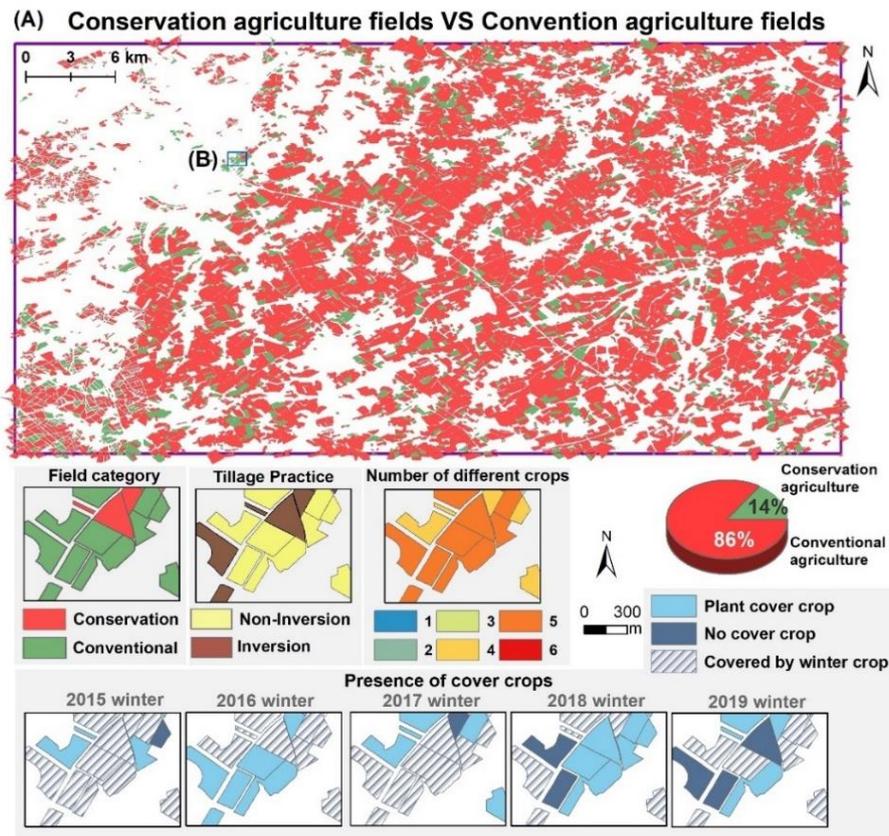


Fig. 5.2. (A) Distribution of conservation and conventional agriculture fields; for zoomed region (B) Field category, tillage practice, number of different crops and annual presence of cover crops are shown (Source: Chapter 3).

The carbon input from plant residues is consists of two components: the carbon input from main crops and the carbon input from cover crops (CC).

1. Main Crop Carbon Input:

We calculated the carbon input of main crops using the Bolinder equation (Bolinder et al., 2007), an allometric equation that relates crop yield to above- and belowground plant biomass to estimate soil C input. Main crop information for each field was obtained from the

LPIS. Provincial yield data were sourced from the Belgian Statistical Office (<https://statbel.fgov.be/>). Detailed calculations see Zhou et al., (2024). However, due to the diversity of crops in the Walloon region and the lack of yield information for about 10% of the crops, we used empirical values from the literature for these crop types (Hendriks et al., 2023). Based on crop calendar information (<https://www.dwd.de/>), we distributed the C input as follows: 50% in the harvest month and 50% equally distributed across the three months preceding harvest.

2. Cover Crop Carbon Input:

We determined the carbon input from cover crops using the annual CC map (Fig. 5.2) identified by Zhou et al. (2024) and long-term LAI data from satellite dataset. For fields where CC was planted in a given year, we obtained the LAI value within the growing season of the CC. LAI was calculated using Landsat 7, 8, and Sentinel-2 data on GEE (Eq.5.4, 5.5; Boegh et al., 2002). We then estimated above-ground biomass (AGB) using a linear model (Eq.5.6) established by Zhou et al. (2024), which relates LAI values to measured AGB at the Lonzée site. The maximum AGB value within the CC growth interval was used as the final AGB value incorporated into the soil. Below-ground biomass (BGB) of CC was calculated using allometric parameters of mustard (C in stem: C in root & extra-root material= 0.537: 0.272), as mustard is one of the most common cover crops (Seitz et al., 2023).

$$EVI = 2.5 * ((NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1)) \quad (5.4)$$

$$LAI = 3.618 \times EVI - 0.118 \quad (5.5)$$

$$AGB = 0.27 \times LAI - 0.19 \quad (5.6)$$

When the EVI values of some fields are missing, the average carbon input value of CC, $1.4 \text{ t} \cdot \text{ha}^{-1} \cdot \text{yr}^{-1}$, is used as a replacement.

We then combined the carbon input data for year 2015-2018 with the soil cover data. We replicated this four-year rotation pattern to reach the end of the simulation (2024).

5.2.3.4 C input from manure

The basic principle for calculating manure application involves obtaining the number of different livestock types in each province from the Belgian Statistical Office (<https://statbel.fgov.be/>). The total manure carbon (C) content for each province is then calculated based on the average annual manure production per type of livestock, the management systems of manure (solid, liquid, etc.), and the carbon content ratio for management system (Chapter 4). It is assumed that the total manure is evenly distributed across all utilized agricultural area (UUA) within each province. The average annual C application from manure for each province from 2015 to 2022, calculated using this method, is as follows (Table 5.2).

Table 5.2. Utilized agricultural area (UUA), agricultural land proportion and annual input of farmyard manure (FYM, t C ha⁻¹) by province

Year	Brabant Wallon	Namur	Liège	Hainaut
UUA (ha)	64,491	210,253	144,973	159,529
Cropland	86%	62%	45%	75%
grassland	13%	37%	54%	25%
FYM 2015	0.39	0.70	0.95	0.78
FYM 2016	0.42	0.73	0.93	0.76
FYM 2017	0.37	0.73	0.88	0.72
FYM 2018	0.39	0.72	0.87	0.72
FYM 2019	0.36	0.71	0.88	0.73
FYM 2020	0.36	0.71	0.87	0.72
FYM 2021	0.35	0.71	0.85	0.70
FYM 2022	0.36	0.70	0.83	0.69

5.2.4 Different spatial scales for the model runs

First, we performed simulations on a field-by-field basis. We parametrized individual RothC runs for 10,102 fields, ran the models from January 2015 to December 2024, and calculated the changes in SOC stock for individual fields and the total C changes for the entire study region. Then, we aggregated the boundary conditions into different spatial stratifications:

1. Pedoclimate zones at the EU scale
2. Soil Associations at the national level

The distribution of these two layers in the study area is shown in Fig. 5.3. Tóth et al. (2016) combined the Reference Soil Group (WRB 1998) with eight climatic groups (Boreal to Sub-Boreal, Atlantic, Sub-oceanic, Northern sub-continental, Southern sub-continental, Mediterranean semi-arid, Mediterranean, Temperate mountainous) in Europe to divide Europe into 133 different pedoclimatic zones. Under this classification, the entire Belgium falls into the Atlantic climate zone, and our study area covers four different soil types.

The Soil Associations Map for Belgium is provided by JRC, European Soil Data Centre (ESDAC) (<https://esdac.jrc.ec.europa.eu/content/carte-des-sols-de-la-belgique-associations-de-sols-pedologie>) and drawn by R. Marechal & R. Tavernier (1970). It divides Belgian soils into 62 subtypes, 16 of which are present in the study area, and 15 of these subtypes cover our fields.

We took the initial SOC stock values of all fields in the sub-zone and weighted them according to the area to obtain the regional SOC stock base value. The boundary condition for the region was calculated by averaging the climate data and carbon inputs (from plants and from manure). Subsequently, the RothC model was run for each sub-region.

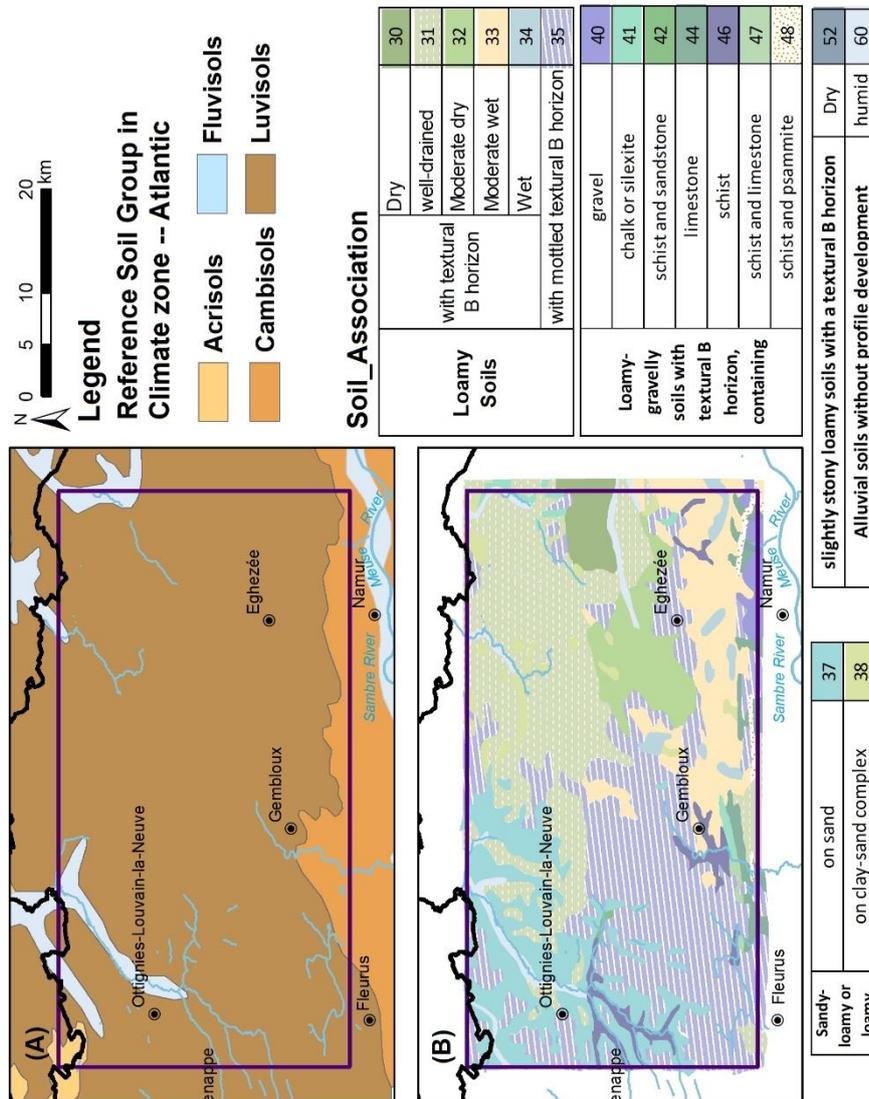


Fig. 5.3. (A) European Pedoclimate zones (B) Belgian Soil Association groups

5.2.5 Model runs for different Management Scenarios

To calculate the carbon sequestration potential of different carbon farming practices, we consider the results of model simulations per field using all measured data based on the crops identified in the LPIS for conventional fields in the previous section as Business-as-usual (BAU). We exclude the fields under conservation agriculture as

identified in Fig. 5.2. This is in line with the current discussions on baselines for carbon farming practices and allows a fair attribution of carbon credits to early adopters of conservation agriculture (Rosinger et al., 2023).

We then set three management scenarios: Scenario 1: Reduced Tillage, Scenario 2: Long Duration Cover Crop, and Scenario 3: Combined Reduced Tillage and Long Duration Cover Crop.

For Scenario 1: Reduced Tillage, we only need to change the fields' tillage information to reduced tillage and set the rate-modifying factor for tillage (d value; Eq. 5.1) to 0.93 when running the model.

For Scenario 2: Long duration Cover crop is more complicated. First, we summarized the existing CC planting situation of all fields (Table 5.3). In general, about half of the fields plant winter crops each year, and about 2/3 (66%-72%) of the remaining fields plant CC.

Table 5.3. Ratio of Planting Cover Crop in Study Area.
(Total fields number: 10,102; CC: cover crop)

	2015	2016	2017	2018
Fields with a winter cereal (a)	5,288	5,005	4,749	4,894
Fields with the potential to grow CC (b)	4,814	5,097	5,353	5,208
Fields planted with a CC (c)	3,208	3,446	3,852	3,431
Ratio of fields planted with a CC/ fields with the potential to grow CC (c/b)	67%	68%	72%	66%
Ratio of fields have potential to plant additional CC/ all fields ((b-c)/(a+b))	16%	16%	15%	18%

Therefore, Scenario 2 involves the following two main steps:

1. Set all fields with the potential to plant additional cover crops as planted with a CC. First, add winter cover crops to the remaining fields without winter crops or cover crop. The carbon input for these

newly added CC is based on the average CC carbon input from other years. If cover crops were not planted in all years, using the average CC carbon input value of all fields ($1.4 \text{ t} \cdot \text{ha}^{-1} \cdot \text{yr}^{-1}$).

2. Set all CC as long-term CC. This involves adjusting the cover rate-modifying factors (value c ; Eq. 5.1) for the period between the two main crops. It is assumed that CC is planted immediately after the harvest of the last main crop, and that the field remains covered from one month after planting CC until the next main crop is planted. The c value during this period is all modified to 0.6.

Scenario 3 is a combination of the two scenarios described above: planting cover crops where possible and applying reduced tillage to every field.

For the running under stratifications, running under Scenario 1 involves modifying the tillage practice to Reduced Tillage. For Scenario 2, since it is difficult to determine the CC coverage for each sub-region, we assume that the fields with potential to grow CC are covered with CC for half of the time over the four-year period. The average CC input is 1.4, with the annual CC input being 0.7. The soil is set as covered from October to April.

5.3 Results and Discussion

5.3.1 The effect of spatial scale on SOC dynamics

We mapped the distribution of SOC stock changes between 2015 and 2024 for the aggregation levels (Fig. 5.4). The single field change maps displayed a more diverse SOC change pattern, with change values ranging from -10.5 to $6.5 \text{ t} \cdot \text{ha}^{-1}$. However, when aggregated, the changes became smaller, ranging from -1.5 to $1.5 \text{ t} \cdot \text{ha}^{-1}$.

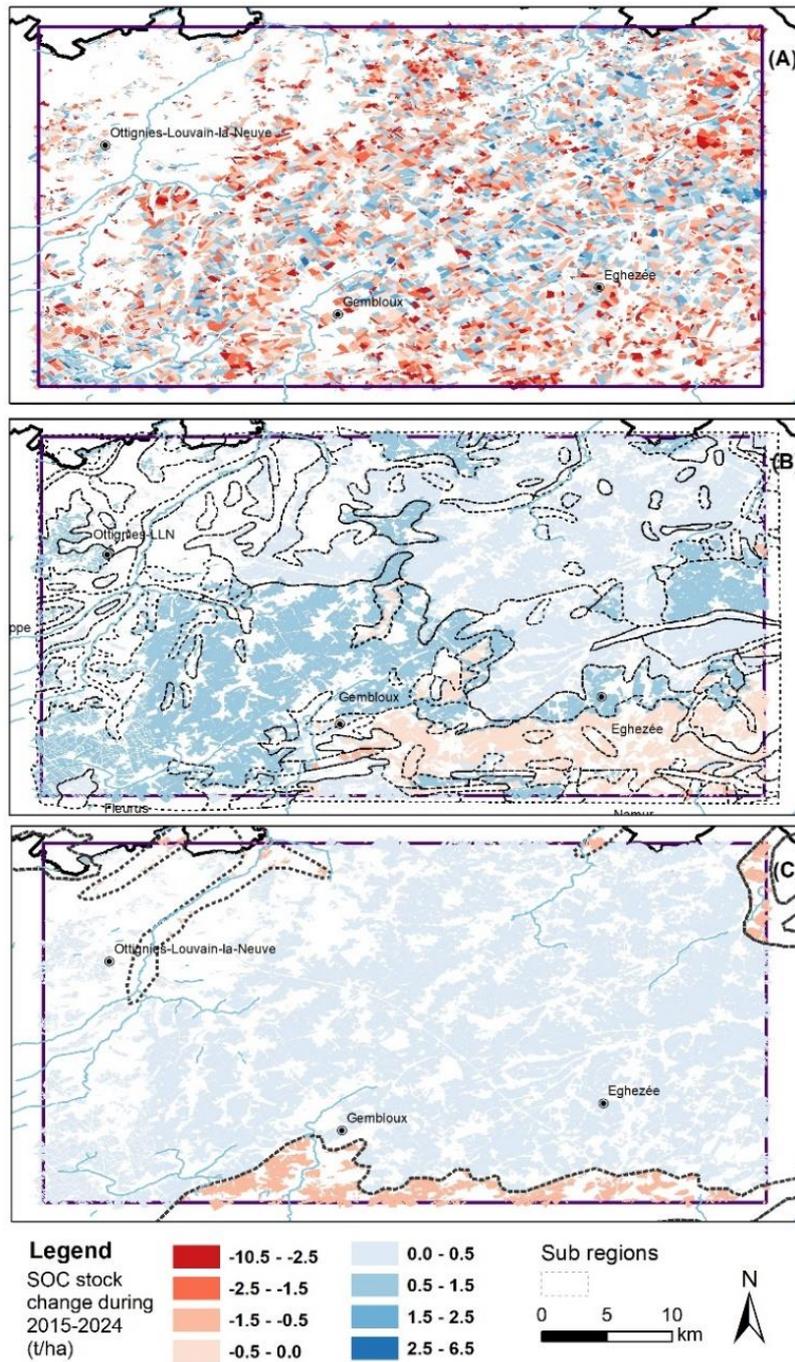


Fig. 5.4. SOC stock changes (t ha⁻¹) during 2015-2024 for model runs at different scales (A) Per field (B) Belgian Soil association groups and (C) European Pedoclimate zones (see Fig. 5.1 for details)

Based on the area-weighted average, we assessed the changes in SOC stock for 50,655.52 ha of cropland in the study area from 2015 to 2024 (Table 5.4). The SOC stock changes based on individual field data shows an average SOC stock loss of $0.38 \text{ t}\cdot\text{ha}^{-1}$. In contrast, calculations for soil associations and soil climate groups show the opposite trend, with increases in soil carbon of $0.31 \text{ t}\cdot\text{ha}^{-1}$ and $0.15 \text{ t}\cdot\text{ha}^{-1}$, respectively. (The detailed SOC stock changes for each subgroup are shown in Tables A.4.1 and A.4.2)

Table 5.4. SOC stock changes during 2015-2024 for 50,655.52 ha of agricultural fields in the Walloon region of Belgium (Here the SOC stock value is the average value of 12 months in the year).

	SOC stock in 2015 ($\text{t}\cdot\text{ha}^{-1}$)	SOC stock in 2024 ($\text{t}\cdot\text{ha}^{-1}$)	Delta ($\text{t}\cdot\text{ha}^{-1}$)
Per fields	46.61	46.23	-0.38
Soil association	46.55	46.86	0.31
Pedoclimate	46.54	46.69	0.15

5.3.1.1 SOC stock changes at field level

In general, the results of independent simulations for each field based on actual management data are more accurate and detailed, which can highlight the differences between fields. Adjacent fields may exhibit completely opposite trends due to differences in management practices.

The map of changes in individual fields (Fig. 5.4 (A)) shows that SOC decreases (colored in red) are more common resulting in an average C loss of $0.038 \text{ t C ha}^{-1} \text{ yr}^{-1}$ (Table 5.4). First, we compared the changes in Fig. 5.4 (A) with the initial SOC stock values shown in Fig. 5.1 (B) and found that the areas with decreasing SOC stock correspond to fields with high initial SOC stock. We calculated the correlation coefficient between the SOC stock change and the initial

SOC stock value, resulting in $r = -0.55$. This relationship may indicate a trend towards the theoretical steady state of soil (Goidts, 2009).

For fields that have consistently remained cropland, the factors influencing SOC stock changes include differences in carbon decomposition rates and variations in external carbon inputs (Jordon et al., 2022). At the field scale, there are no notable climatic differences between adjacent fields, so changes in SOC stock are mainly driven by differences in carbon input and soil coverage. We examined fields with evident degradation and identified several key factors contributing to the continuous decline in SOC stock:

1. High frequency of root crop planting: Root crops often result in lower residual carbon input compared to cereal crops.
2. Prolonged soil exposure during winter: Long periods of bare soil, with no cover crops planted after the main crop harvest, negatively impact SOC levels.
3. Long-Term single winter cropping without cover crops: Continuous planting of winter crops without the use of cover crops can degrade soil structure and deplete nutrients.

The reasons and literature evidence that these management measures will cause C stock to decline have been discussed in Chapter 4. Conservation agriculture practices showed in Fig. 5.2 can address these factors these issues and improve soil health.

By combining the field-level prediction results with agricultural systems (conventional or conservation), we created a density distribution plot (Fig. 5.5) that illustrates SOC changes over the ten-year period from 2015 to 2024. The black lines in the figure represent the uncertainty values derived from bias in the long-term experiment simulations (Chapter 4), indicating that around 70% of fields had no evident SOC changes over the past decade. Nevertheless, the current CA practices have yielded positive outcomes, with a higher proportion (8.5%) of fields showing an increase in SOC stock. However, the implementation of CA remains insufficient to fully eliminate soil

carbon degradation, as the mean SOC change value is still negative. Similar findings were reported in long-term experimental sites in Switzerland, where measures designed to mitigate the environmental impacts of agricultural systems, implemented by farmers to receive subsidies, were not sufficient to preserve or increase SOC (Keel et al., 2019; Guillaume et al., 2021).

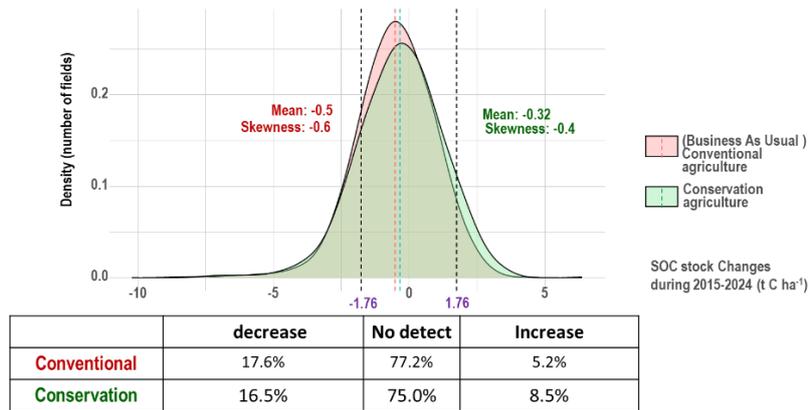


Fig. 5.5. Density plot of SOC stock changes in conventional and conservation agriculture systems from 2015 to 2024.

For the conventional fields, representing the BAU scenario, potential C sequestration through the implementation of sustainable or conservation management practices are detailed in Section 3.2.

5.3.1.2 SOC stock changes after aggregation

After aggregating the model simulations, it becomes easier to assess the impact of climate factors and soil types on SOC stock changes at a larger scale.

According to the European classification of Pedoclimate zones, the entire study area falls within one climate type—Atlantic. Consequently, variations in SOC stock are primarily influenced by soil type. Specifically, SOC stocks decreased in soils classified as Acrisols and Fluvisols, while they increased in soils categorized as Cambisols and Luvisols. In our study area, only four fields were classified as Acrisols, which were not representative enough to be

included in our analysis. The reason that Fluvisols and some soil association groups in humid environments (Figure 5.3, Groups 33 and 34) show a decrease in SOC is due to two main reasons. First, they have a high initial SOC stock and therefore a higher loss amount, as can be inferred from Equation 5.1. Second, over the long term, strong drainage events will increase the leaching of dissolved organic carbon in agricultural soils, leading to a decrease in SOC stock (Garnier et al., 2022; Conchedda & Tubiello, 2020).

In addition, as shown in Fig. 5.4 and Table 5.4, the simulation level has a substantial influence on the results. Many studies simulate SOC stock based on HLUs (Wiesmeier et al., 2016; van Wesemael et al., 2010; Morais et al., 2019; Lugato et al., 2013), which offers the advantage of requiring fewer calculations, as only one simulation is needed per unit. For instance, in our study, we ran simulations 4 different pedoclimatic zones and 16 times for soil association groups. In contrast, RothC required 10,102 simulations for individual fields, which is very time-consuming and labor-intensive. However, our results show that field-scale simulations can reveal more detailed differences, whereas aggregation tends to lose many variations and can result in an entirely opposite trend in overall carbon change.

In addition, one of the drawbacks of RothC is its computational intensity in large-scale simulations because it operates at the field scale. In contrast, landscape or regional-scale earth system models can cover broader areas more efficiently. Le Noë et al. (2023) conducted a comprehensive review of approximately 250 models, identifying 40 that can simulate at large scales. For example, our study area is covered by eight 0.125° ORCHIDEE grids, allowing ORCHIDEE to run the model only once. However, it's important to note that large-scale models also estimate average climate, land use, and soil data within each grid, which can introduce errors similar to those found in aggregation.

5.3.2 SOC Stock changes under different Scenarios

Fig. 5.6 illustrates the density distribution of SOC stock changes in conventional agricultural fields between 2015 and 2024 under different management scenarios for each field. Compared with the BAU scenario, both reduced tillage (RT) and the long-term cover crops (CC) can effectively enhance SOC stock, consistent with the conclusions of Castaldi et al. (2024). Under BAU, only 5.2 % of the fields showed clear increase in SOC stock, while 17% experienced a decrease. However, with RT and CC, the percentage of fields showing SOC increases rises to 11.8% and 32.7%, respectively. By implementing both management practices simultaneously, half fields could have a notable SOC stock increase by the end of 2024.

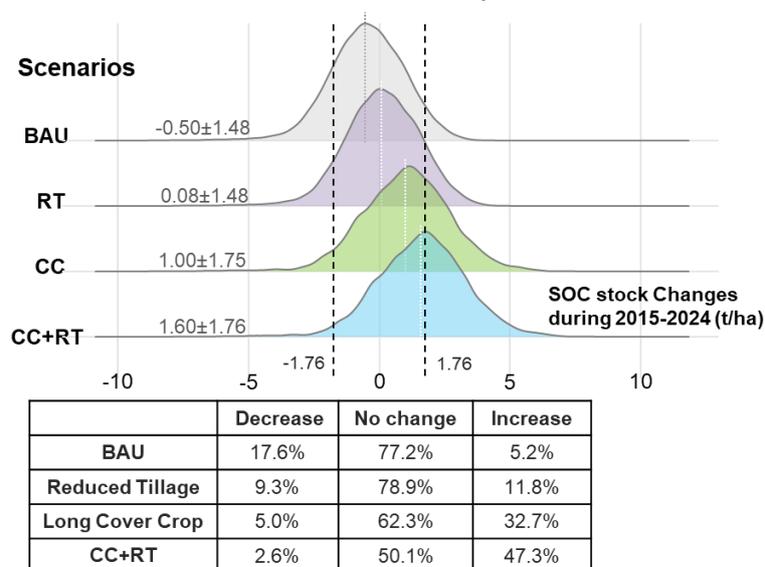


Fig. 5.6. Density plot of SOC stock changes per field under different management scenarios from 2015 to 2024, dash line represents the uncertainty value obtained from section 4.3.2 (BAU: Business-as-usual; RT: reduced tillage; CC: Long duration cover crop).

Table 5.5 quantifies the change in SOC stock and CO₂ emissions under different scenarios, based on a weighted average of area per field. Under BAU, current carbon stocks cannot be sustained, leading to an average loss of -0.39 t·ha⁻¹ of carbon over ten years for

44,004.19 ha cropland. However, implementing conservation measures can reverse this carbon loss, shifting the soil from a carbon source to a carbon sink, with SOC stock increasing by 0.26 to 1.8 t·ha⁻¹ over the same period. Compared to BAU, the combined implementation of CC and RT can achieve a carbon sequestration potential of 0.22 t·ha⁻¹·yr⁻¹, resulting in an additional accumulation of 350,000 t of CO₂. Moreover, this scenario results in an annual carbon stocks increase rate of 3.8‰, which is an encouraging sign towards meeting the goal of the 4 per 1000 initiative.

Table 5.5. Soil organic carbon (SOC) stock changes during 2015-2024 for 44,004.19 ha of conventional agricultural fields in the Walloon region of Belgium, calculated based on per field (Here the SOC stock value is the average value of 12 months in the year).

	SOC stock 2015 t·ha ⁻¹	SOC stock 2024 t·ha ⁻¹	Δ C compared to 2015 t·ha ⁻¹	Sequestration Rate of management t·ha ⁻¹ ·yr ⁻¹	Δ CO ₂ Mt
BAU	46.42	46.023	-0.39		
RT		46.68	0.26	0.07	0.11
CC		47.56	1.14	0.15	0.25
CC+RT		48.22	1.80	0.22	0.35

The average carbon sequestration potential of CC (0.15 t C ha⁻¹ yr⁻¹) obtained in this study is relatively low compared to other studies. Poeplau & Don (2015) estimated a global carbon sequestration potential for cover crops of 0.32 t C ha⁻¹ yr⁻¹, while Porwollik et al., (2022) suggest a range of 0.48-0.52 t C ha⁻¹ yr⁻¹. This is because cover crops in Belgium already have good initial conditions. Specifically, maintaining minimum soil cover during winter, particularly in erosion-prone areas, is a requirement for farmers to receive payments under the Common Agricultural Policy (2023-2027; <https://agriculture.wallonie.be/accueil>). As shown in Table 5.3, only 15%-18% of the total fields have potential for additional planting CC, much lower than values reported in other studies (Global: 25%,

Porwollik et al., 2022; Germany: 30%, Seitz et al., 2023; Denmark: 20–25%, Taghizadeh-Toosi & Olesen, 2016). This highlights the importance of considering up-to-date implementation of practices when estimating the carbon sequestration potential of management measures in a specific area.

It is also important to highlight that our estimates for the RT scenario assume a positive impact of reduced tillage on SOC stocks within the upper 30 cm by reducing the fragmentation of soil aggregates, which slows down the decomposition rate of soil carbon (Wang et al., 2019). We applied the rate-modifying factor directly proposed by Jordon & Smith (2022). However, recent meta-analyses suggest that in temperate environments, reduced tillage may lead to a redistribution of SOC within the soil profile, with increases in SOC stocks in the 0-10/15 cm layer but decreases in the 15/20-30 cm layer (Krauss et al., 2022). Additionally, a meta-analysis by Virto et al. (2011) indicates that carbon inputs difference from crops is the only factor significantly and positively related to SOC stock differences between no-till and inversion tillage systems. Our model does not include a detailed tillage module, nor does it simulate the complex changes in aggregate structure, fragmentation, and SOC distribution across the soil profile, which would require a more comprehensive algorithm of the relevant physical and chemical processes.

5.3.3 The effect of the choice for a baseline on calculating C sequestration.

Here we illustrate the effect of different aggregation methods for the baseline on carbon sequestration by cover crops (Fig. 5.7). We show the RothC model outputs for some examples of individual fields (Fig. 5.7 A and B). We also present examples of RothC runs using average values for the entire region (Fig. 5.7 C) and for two soil associations according to the Belgian soil map (Fig. 5.7 D). The blue

line represents BAU, and the green line represents long cover crops (CC). The difference in SOC stock between the treatment and the baseline at the end of the model run represents the simulated carbon sequestration. For different examples of fields, the simulated sequestration is quite different. After checking the detailed management practices of these fields, it was found that the fields in Fig. 5.7(A) do not plant cover crops in the rotation, while the fields in Fig. 5.7(B) grow cover crops but plough them into the soil after only a few months, typically in winter. The simulated increase in SOC stock in these fields is related to the longer growing season of the cover crops (Fig. 5.7 B). These two cases demonstrate the need for high resolution boundary conditions of the models or measuring carbon stocks over time, as is the case for current carbon crediting or payment schemes (Rosinger et al., 2023).

Simulations with boundary conditions averaged over entire areas (see section 5.3.1.2) are not able to capture the impact of soil climate conditions (especially soil texture) and initial SOC stock on soil sequestration potential is ignored (Gutierrez et al., 2023; Wiesmeier et al., 2019; Rosinger et al., 2023). When sub-regions (Soil association groups) are used, the historical carbon management practices are often disregarded. Moreover, early adoption can lead to reduced carbon sequestration potential, as seen in Case 2 and Case 4 (in Fig. 5.7 (A) and (B)). This results in lower returns from pioneering farmers who initially implemented improved management practices, raising concerns about the fairness of carbon farming schemes (Rosinger et al., 2023).

Therefore, it still needs to be discussed whether payments need to be related to specific practices that promote soil carbon sequestration (activity base), rather than to actual SOC changes quantified by SOC measurements (results based). This view is also supported by Guillaume et al. (2021)

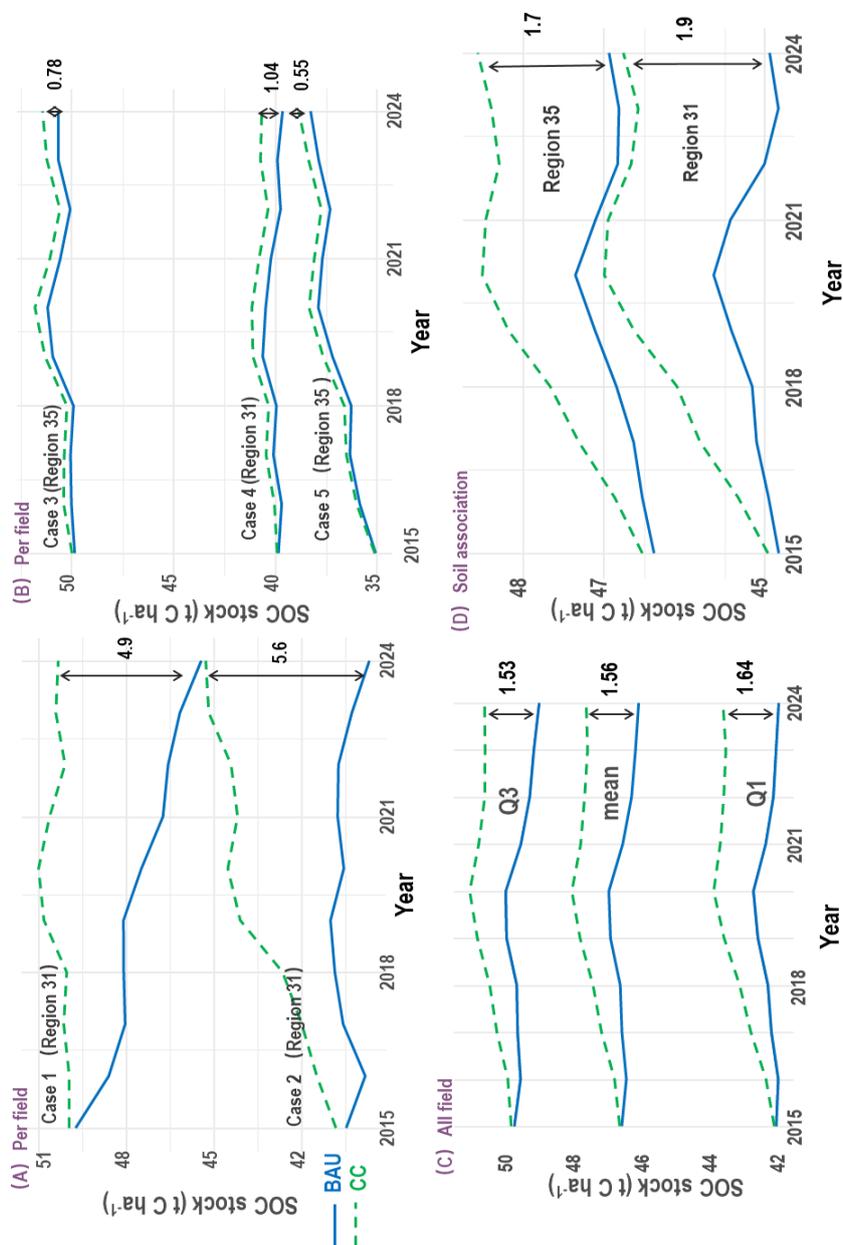


Fig. 5.7. Compare SOC stock changes (t·ha⁻¹) between Business-as-Usual and Long-term Cover crop Scenarios during 2015-2024 for (A) two example fields (B) Three example fields (C) Averaged value for entire region and (D) Two soil associations. (Region 31: Loamy Soils with textural B horizon (well-drained); Region 35: Loamy Soils with mottled textural B horizon)

5.3.4 Limitations

A major limitation of regional SOC change predictions is the lack of precise, field-specific data on both manure application and crop yield. The amount of manure applied can vary widely between fields, which can substantially impact SOC stock (Li et al., 2021; Gross & Glaser, 2021; Maillard & Angers, 2013). In this study, we had to rely on provincial-level averages due to the unavailability of field-specific data, which introduces uncertainty into the predictions. One feasible solution is to improve crop yield estimates using crop models such as SAFYE-CO2 (Pique et al., 2020), which enables spatialized estimation of yield. As this model is driven by high-resolution satellite data, it provides the potential for more accurate, field-specific yield predictions. Additionally, the climate data used in this study is not strictly remote sensing data but rather reanalysis data. To improve accuracy, future studies could consider using more up-to-date remote sensing products to obtain high-resolution temperature and precipitation data.

Another obvious limitation is that we used data from the main crops between 2015 and 2018 as a four-year rotation. In reality, the rotation of different fields is not fixed and may vary from three to six years. There is even no standard crop rotation; instead, farmers decide on which crops to grow based on field conditions, weather patterns, and economic considerations. In some cases, they may even introduce temporary grasslands into the rotation. This causes our replication of a four-year rotation to deviate from actual practices.

To enhance the analysis, we could extend the data collection period for main crops, identify specific rotation patterns, and replicate those rotations. However, this approach would require more comprehensive field-level census data. Alternatively, we could introduce the "sequence analysis" method, commonly used in land use

change studies, to evaluate crop rotations. This method would allow us to assess additional indicators such as the frequency, sequence, and duration of rapid changes in crop patterns (Guillaume et al., 2021; Watson et al., 2013).

When simulating SOC changes under different scenarios, one of the shortcomings of our study is that it didn't consider the role of climate change. In fact, numerous studies have shown that rising temperatures will reduce SOC stock (Kirschbaum, 1995, 2000). Wiltshire and Beckage (2023) used static climate conditions from 1991 to 2021 and the climate change scenario RCP4.5 to simulate SOC changes under different land management practices from 2022 to 2099. Their results showed that rising temperatures would cause SOC stocks to decrease by 9.1% to 19.9%, leading to a net SOC loss even under many regenerative farming scenarios. The RothC model does not have a plant growth module. Changes in climate conditions only change the decomposition rate constant, and there is no mechanism to measure how external carbon input varies with climate. A feasible solution is to calculate carbon input using net primary production (NPP) values derived from precipitation and climate data, allowing the NPP to be scaled accordingly. Poeplau & Dechow (2023) simulated the dynamics of SOC storage using RothC in global agricultural surface topsoil (0-30 cm) from 1919 to 2018. They found that, over 100 years, a global temperature rise of 1.03 °C could result in a SOC loss of $2.5 \pm 2.3 \text{ t}\cdot\text{ha}^{-1}$ if NPP remains unchanged. When accounting for the impact of climate change on NPP, the SOC loss averages $1.6 \pm 3.4 \text{ t}\cdot\text{ha}^{-1}$. It can be observed that the impacts on climate change are often considered over much longer time scales. Annual changes in SOC due to changing climate ($0.016\text{-}0.025 \text{ t}\cdot\text{ha}^{-1}\cdot\text{yr}^{-1}$, Poeplau & Dechow, 2023) are far smaller than SOC changes resulting from management practices (0.07 to $0.22 \text{ t}\cdot\text{ha}^{-1}\cdot\text{yr}^{-1}$) as obtained from our study. However, it is worth trying to use earth system models that

includes ecosystem processes for simulation in subsequent studies, considering the dual impact of future climate change and changes in management measures.

5.4 Conclusions

In our study area, we simulated changes in SOC stock across 10,102 fields. At the individual field level, SOC stock showed a decreasing trend, with a mean SOC stock loss of $-0.38 \text{ t} \cdot \text{ha}^{-1}$ over 50,655.52 ha of cropland within ten years. Although some pioneering conservation agriculture measures have been implemented over the past decade, these practices have not yet altered the trend of carbon loss. However, if reduced tillage and cover crop management measures are thoroughly and effectively implemented, even conventional fields could achieve an annual increase in carbon storage of 3.8 per thousand. When aggregating the results from individual fields to a soil association or a region, it is important to note that the predicted outcomes may carry a considerable risk of inaccuracy.

Chapter 6.

Summary and Conclusions

6.1. Main Findings

Overall, this thesis has achieved a comprehensive simulation process for the spatio-temporal variability and sequestration potential of SOC in cropland. We mapped the spatial distribution of SOC at the base period using digital soil mapping (Chapter 2), identified regional management practices using remote sensing data (Chapter 3), simulated the temporal dynamics of SOC with data from a flux tower station in cropland (Lonzée), and used remote sensing data to replace all in-situ boundary conditions (Chapter 4). Ultimately, based on the validated RothC model, base period SOC map, and regional climate and management boundary conditions, we simulated the spatio-temporal changes in regional SOC and established management scenarios to discuss the carbon sequestration potential under different management scenarios (Chapter 5).

The key findings of this thesis are as follows:

Chapter 2: When using digital soil mapping methods, it is highly effective to include covariates related to human activities. We defined a covariate (OC input) to measure the carbon input from the history crop cultivation of cropland. Overall, by using the GBM model, we successfully predicted the spatial distribution of base period SOC, with an R^2 value of 0.77. Analysis using Empirical Mode Decomposition revealed that climate and topography are the primary factors determining large-scale SOC distribution, while differences between fields mainly result from variations in organic carbon input

from crop residues. This underscores the critical importance of accurate management practice information for fine-scale simulation of SOC spatial variation in cropland fields.

Chapter 3: It is feasible to predict large-scale agricultural management practices using RS methods. By employing long-term optical and radar satellite data along with precipitation data, we developed a conservation agriculture classification model (overall accuracy 92%) for the Walloon region of Belgium. We found that the adoption rate of conservation agriculture in the Hesbaye region is 15.5%, higher than current public statistics. We also created some model covariates based on the three principles of conservation agriculture, simulating field management practices, mainly including crop rotations, presence and duration of cover crops and tillage practice. Our methods for simulating cover crops and tillage practices, based on long-term RS data, can provide monthly or annual management information for process-based models.

Chapter 4: An important basis for accurately simulating long-term SOC changes using the RothC model is the precise estimation of C input. Simulation of SOC changes at the Lonzée station from 2007 to 2017 showed that SOC fluctuated due to the application of farmyard manure and crop residue inputs, with almost no clear change in the initial and final values. We provided two different ways to get boundary conditions: in-situ or RS, and found a bias of $1.76 \text{ t C} \cdot \text{ha}^{-1}$ when replacing site-measured boundary conditions with remote sensing data. The largest bias was due to C input from main crops and cover crops. The boundary conditions calculated based on remote sensing data are fundamental for extending process-based models from point scale to landscape scales.

Chapter 5: From 2015 to 2024, SOC stocks in the study area have acted as a carbon source, reflecting a net carbon loss. Over the 10-year period, the SOC stock showed a decreasing trend, with a mean

loss of $0.38 \text{ t} \cdot \text{ha}^{-1}$. If reduced tillage and long-term cover crop management measures are effectively and comprehensively implemented, even conventional fields could achieve a carbon increase rate of 3.8‰, and result in an additional accumulation of 350,000 t of CO₂. However, when aggregating field prediction to a lower resolution (such as pedoclimate zones or soil association groups), the overall region will convert from a carbon source to a carbon sink. This completely opposite trend highlights that while aggregating subregions can save significant computational time, it may introduce great accuracy risks.

6.2. Limitations and Perspectives

First, the distribution of soil samples is crucial as the base period SOC distribution map forms the foundation and starting point of entire work. The accuracy of SOC predictions under digital soil mapping is closely linked to the distribution, density, representativeness, and quality of the original soil samples. An important issue in our study is the uneven distribution of soil samples, which is particularly dense in the northern Walloon region and sparse in the southern regions. Standard cross-validation may yield overly optimistic accuracy assessments due to the high sampling density in certain areas. Future research should consider more appropriate sampling strategies, especially in areas with large spatial heterogeneity, or using spatial cross-validation methods. Additionally, the soil sample collection spanned a considerable time period from 2015 to 2019, and we assume that the SOC map based on this period represents the initial state from 2015, although there is actually a time difference.

Furthermore, our assessment of management practices relies on time series remote sensing data. In this thesis, we have integrated the optimal publicly available optical and radar remote sensing data,

including Sentinel-1 and Sentinel-2, combined with environmental precipitation data, to predict cover crop adoption and tillage practices across the region. The cover crop predictions based solely on remote sensing data can be readily applied to other countries and contexts; however, for the tillage model, although the workflow can also be easily applied in other regions, it's essential to have field-collected tillage information to calibrate the model. Also, it is important to consider that management practices often occur within short time windows, such as tillage, which may occur within a short interval before seedbed preparation. Consequently, data loss due to adverse weather conditions (e.g., cloud cover) can significantly impact the final prediction results. Future research should consider time series data interpolation to fill the gaps. And so far, we can only estimate the application of farmyard manure amount based on regional statistical data, although it is one of the most important factors affecting cropland SOC dynamics. In practice, the amount of manure applied can vary significantly between fields. However, it is not yet possible to effectively model it through remote sensing data. This will introduce uncertainty into regional SOC simulations.

In addition, the simulation of time-series data based on long-term experiment lacks independent time-series validation, which is crucial for reliable prediction. Moreover, two SOC inventory measurements were taken at the Lonzée site, and there were little changes in the start-end SOC levels. Consequently, the simulation of dynamics within this interval lacks sufficient validation. A potential solution would be to include more long-term experimental sites, use spatial data as a substitute for temporal data (Le Noë, 2023), or collect additional time-series measurements.

Lastly, this thesis only simulated SOC changes under different scenarios, without considering the effects of climate change. Climate change can significantly influence SOC decomposition rates and

vegetation growth. Although its impact on SOC is smaller than that of management practices, it remains an important factor (Beillouin et al., 2023). However, the RothC model does not include a module for plant growth, limiting its ability to fully capture these dynamics. To address this, we could adjust carbon inputs by scaling them according to net NPP or crop yield, or alternatively, use Earth system models that incorporate ecosystem processes.

Generally, our work serves as a reference for using process models to investigate large-scale SOC stock spatio-temporal variation. It aids in understanding carbon stocks in the Walloon region of Belgium, assists in monitoring and managing soil resources, helps evaluate the current implementation of conservation agriculture, identifies potential fields or farms for soil carbon sequestration projects, and provides carbon baselines for greenhouse gas emission trading schemes. Our work has created a scalable, accurate, and reliable soil carbon monitoring, reporting, and verification method that meets Tier 3 standards.

Moving forward, future research should consider employing Earth system models that incorporate ecosystem processes, such as the ORCHIDEE model, to enhance the ability to simulate SOC changes under climate changes. Additionally, it would be beneficial to apply the models to various land use types, such as Belgium's forest and grassland systems. Expanding the scope to evaluate greenhouse gas trends across broader scales (national, EU, or even global) under carbon removal projects will further support soil carbon monitoring, reporting, and verification protocols.

Appendix

Appendix 1 (Chapter 2)

Table A.1.1. Organic carbon input from crop residues (OC) and humification coefficient (HC) of different crop types.

Crop type	OC	HC	Type
Grassland	2.90	0.31	all year
Winter barley	2.24	0.31	winter
Spring barley	1.88	0.31	summer
Winter wheat	2.08	0.31	winter
Spring wheat	2.08	0.31	summer
Winter rapeseed	1.95	0.27	winter
Spring rapeseed	1.55	0.25	summer
Maize grain	3.20	0.31	summer
Maize silage	1.25	0.34	summer
Sugar beet	2.30	0.22	winter
Carrot	1.21	0.25	summer
Beans	0.91	0.24	summer
Flax	0.14	0.33	summer
Chicory	1.40	0.3	summer
Potato	1.76	0.23	summer
Other vegetables	1.78	0.26	summer
Cover crops	1.46	0.31	

Table A.1.2. Hyperparameters grid

Hyperparameters	Names in <i>gbm</i> package	Values
Shrinkage	shrinkage	(0.01, 0.05, 0.1)
Depth of interaction	interaction.depth	(3, 5, 7)
fraction of bagging	bag.fraction	(0.6, 0.8, 1)
trees terminal nodes	n.minobsinnode	(6, 8, 10)

Appendix 2 (Chapter 3)

Table A.2.1. Crop calendar of main crop types in Hesbaye region (The frequency during 2015-2020, data source: <https://geoportail.wallonie.be/>). (*represent start from previous year) (F: Frequency %)

Crop type	F	Seeding mm/dd	harvest mm/dd	Type	Source of date
Winter wheat	38.4	10/01*	07/30	winter	dwd.de
Sugar beet	13.6	04/15	10/25	summer	getmeteotrack.com
Potato	11.3	04/15	09/20	summer	dwd.de
Maize silage	6.7	05/01	09/25	summer	dwd.de
Winter barley	6.0	10/01*	07/15	winter	dwd.de
Flax	4.2	04/01	08/10	summer	agro-v.com
Peas	3.6	05/01	08/20	summer	almanac.com
chicory	3.0	03/15	10/31	summer	growveg.com
Maize grain	1.8	05/01	11/05	summer	dwd.de
Winter rapeseed	1.5	8/13*	07/20	winter	dwd.de

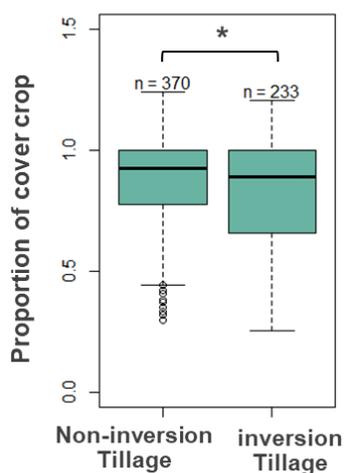


Fig. A.2.1 Boxplot of proportion of cover crop duration in the entire interval under inversion and non-inversion tillage (*: $P \leq 0.05$, significance notation under Wilcoxon test)

Fig. A.2.2 Correlation plot between seven continuous covariates, where a cross denotes the relationship is not significant (Here: N_crop: number of different annual crops; N_cereal: number of cereals; N_CC: number of cover crops; Du_CC: duration of the cover crop; P_BS: bare soil period; P_fallow: fallow period)

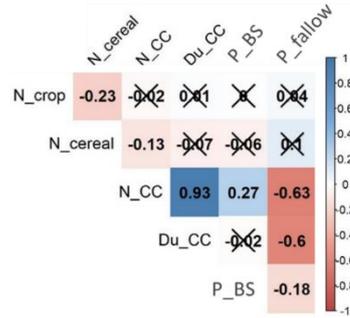
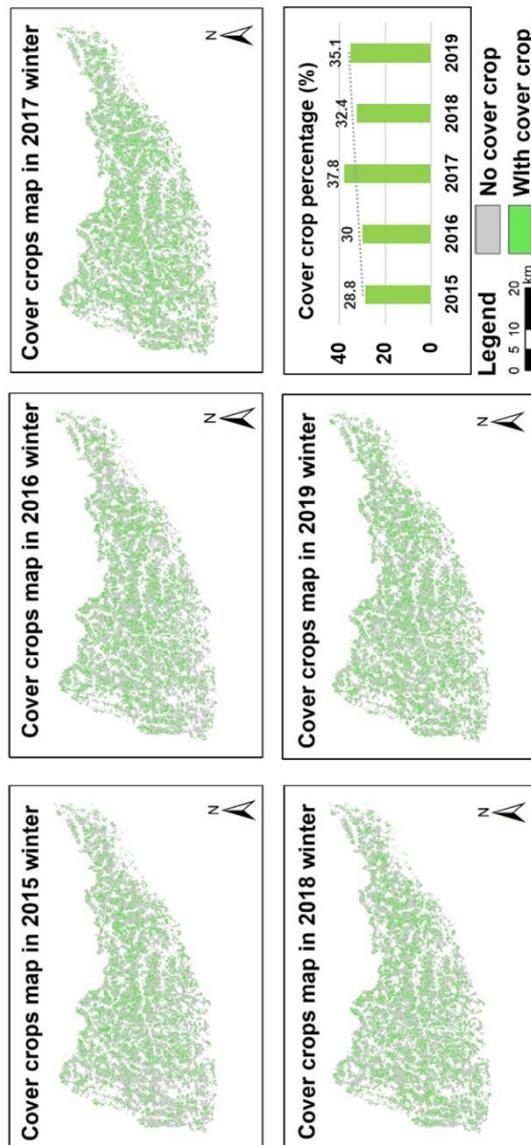


Fig. A.2.3 Cover crops maps during 2015-2020 in Hesbaye region



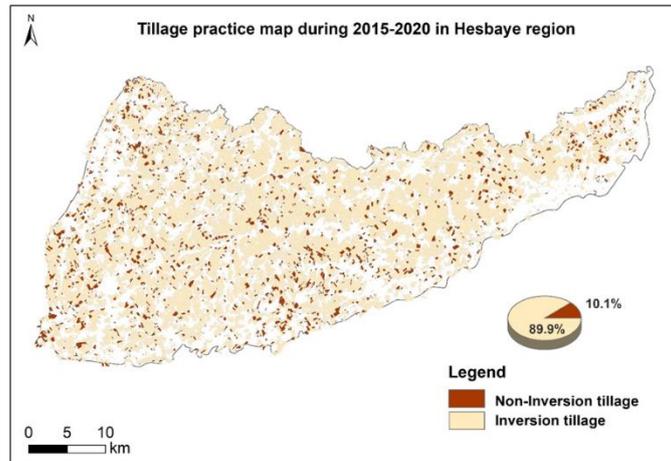


Fig. A.2.4 Tillage practice map during 2015-2020 in Hesbaye region

Appendix 3 (Chapter 4)

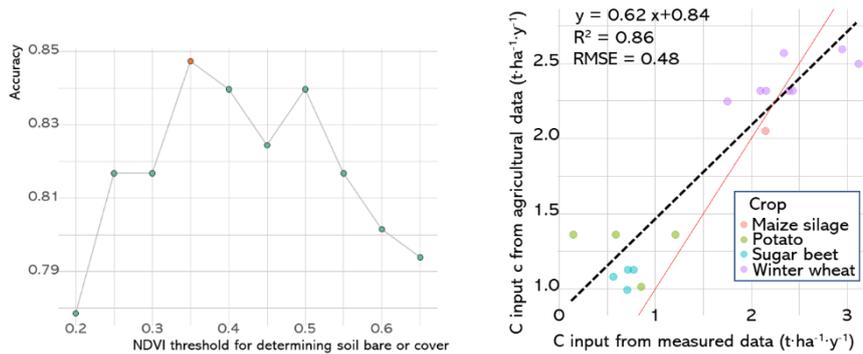


Fig. A.3.1. (left) Overall accuracy of the binary soil cover indicator corresponding to different NDVI thresholds

Fig. A.3.2. (right) Comparison of annual C input calculated from different data source

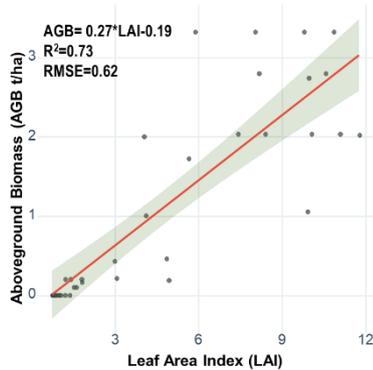


Fig. A.3.3. LAI vs. AGB with regression line and 95% confidence interval

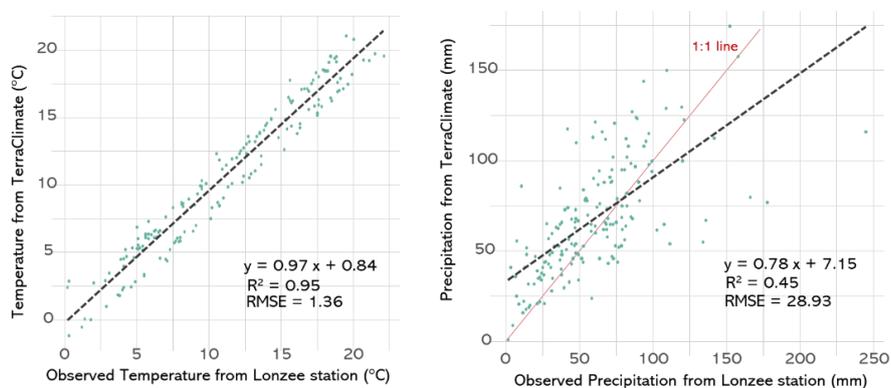


Fig. A.3.4. Comparison of (a) Temperature and (b) Precipitation from Observed and Terraclimate Data.

Table A.3.1. Boundary conditions and their functions in RothC model

	Variable	Units	Function
Climate data	Monthly mean air temperature	°C	Calculate rate modifying factor a
	Monthly rainfall	mm	Calculate rate modifying factor b
	Monthly open pan evaporation	mm	Calculate rate modifying factor b
Soil info	Clay content of the soil	%	Calculate rate modifying factor b Proportion goes to CO ₂ or (HUM+BIO)
	Depth of soil layer sampled	cm	30
	Initial carbon pool	t/ha	
Parameter	DPM/RPM ratio	none	Proportion C goes to DPM and RPM
Management data	Soil cover (0 or 1)	none	Calculate rate modifying factor c
	Monthly input of plant residues	t C ha ⁻¹	As C input
	Monthly input of farmyard manure	t C ha ⁻¹	As C input

Table A.3.2. Estimated manure for different livestock types ($t \cdot head^{-1} \cdot year^{-1}$) (Source: Scarlat et al., 2018)

	Fresh manure t/head/year		Fresh manure t/head/year
calves	2.9	sows	4
bovine	7.3	sheep	0.5
male bovine	9.1	goat	0.5
dairy cows	19.3	broilers	0.04
other cows	9.1	laying hens	0.07
piglets	0.2	other poultry	0.11
other pigs	1.6		

Table A.3.3: the proportion (%) of various management systems for some livestock is offered can also be found in previous work (Source: IPCC, 2019)

	Liquid /Slurry	Solid storage	Pasture/ Range/ Paddock	Daily spread	Others
Dairy Cattle	43	29	26	2	
non-Dairy Cattle	22	26	48	4	
Pig	51				49

Table A.3.4: Parameters of C content ratio of manure (wet weight) (Source: Hendriks et al., 2023)

Type of manure	Carbon/ total wet weight ($t \cdot t^{-1}$)
cow manure (slurry)	0.04
cow manure (solid)	0.08
cow manure (pasture)	0.04
Pig manure (slurry)	0.04
Pig manure (solid)	0.08
Sow manure (slurry)	0.01
Chicken manure	0.21
Other	0.02
Sheep/goat manure	0.09

Table A.3.5. Coefficients for the Bolinder equation (Source: Bolinder 2007; Keel et al., 2017; Dechow et al., 2019; Gan et al., 2009)

<i>Crop</i>	<i>RP</i>	<i>RS</i>	<i>RR</i>	<i>RE</i>	<i>SP</i>	<i>SS</i>
<i>Cereals</i>	0.335	0.482	0.11	0.073	0	0.15
<i>Winter wheat</i>	0.322	0.482	0.118	0.078	0	0.15
<i>Spring wheat</i>	0.322	0.482	0.118	0.078	0	0.15
<i>Winter barley</i>	0.451	0.4	0.09	0.059	0	0.15
<i>Spring barley</i>	0.451	0.4	0.09	0.059	0	0.15
<i>Oats</i>	0.319	0.283	0.241	0.157	0	0.15
<i>Triticale</i>	0.26	0.506	0.142	0.092	0	0.15
<i>Grain</i>	0.219	0.656	0.075	0.05	0	0.15
<i>Maize grain</i>	0.386	0.387	0.138	0.089	0	0.1
<i>Soya beans</i>	0.304	0.455	0.146	0.095	0	0.1
<i>Maize silage</i>	0.772	0	0.138	0.09	0.05	0
<i>potato</i>	0.727	0.232	0.025	0.016	0	1
<i>Sugar beet</i>	0.619	0.353	0.017	0.011	0	0.15
<i>peas</i>	0.217	0.507	0.167	0.109	0	0.1
<i>Rapeseed</i>	0.313	0.383	0.174	0.13	0	0.15
<i>Flax*</i>	0.197	0.661	0.086	0.056	0	0.15
<i>Mustard</i>	0.209	0.537	0.154	0.118	0	1

Scaling factor, SF, for belowground C inputs (Gale & Grigal, 1987; Keel et al., 2017; Dechow et al., 2019):

$$SF = \frac{1 - \beta^{depth}}{1 - \beta^{depthAllomEq}} \quad (A.3.1)$$

Where β is a biome-specific parameter (0.961 for crops, 0.943 for grassland (Jackson et al., 1996)). $depthAllomEq$ is the depth for the allometric equation (for Bolinder, 40 cm for crops, 20 cm for grassland)

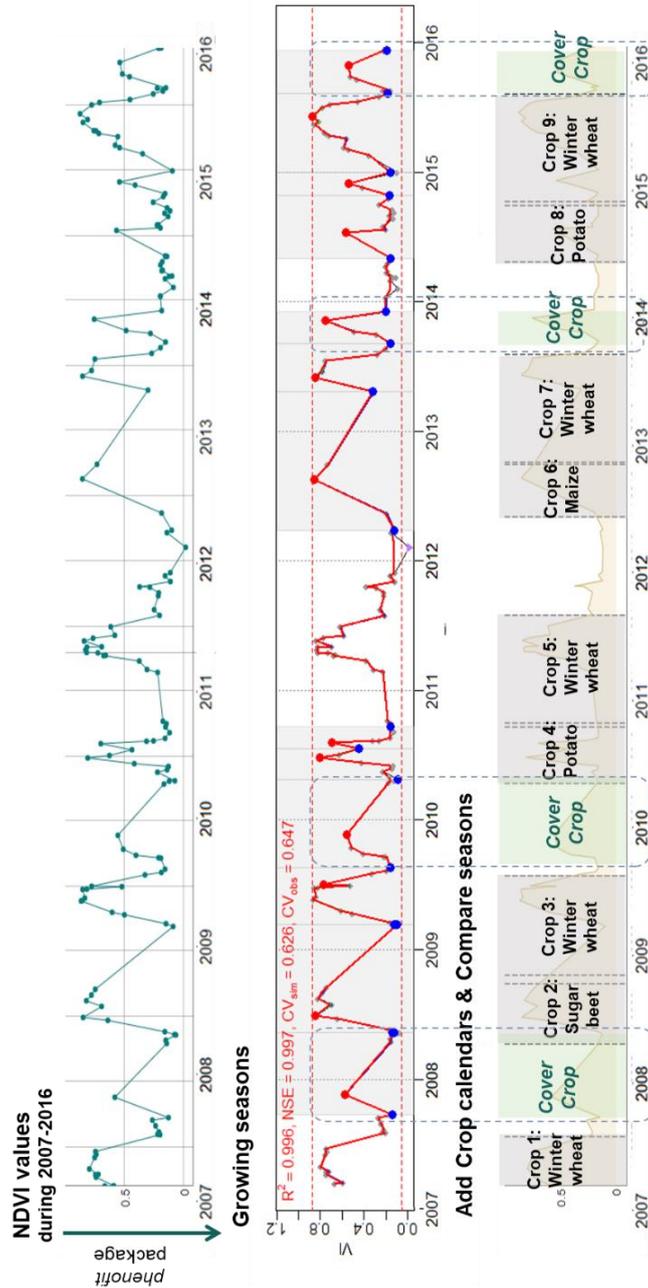


Figure A.3.6. Identifying cover crop of Lonzée station for period 2007-2016
 (*) *For the cover crop in the winter of 2009: There should be an NDVI valley at the end of December (based on measured AGB information), but the earliest available remote sensing (RS) data is in April 2010, so the duration of CC is overestimated

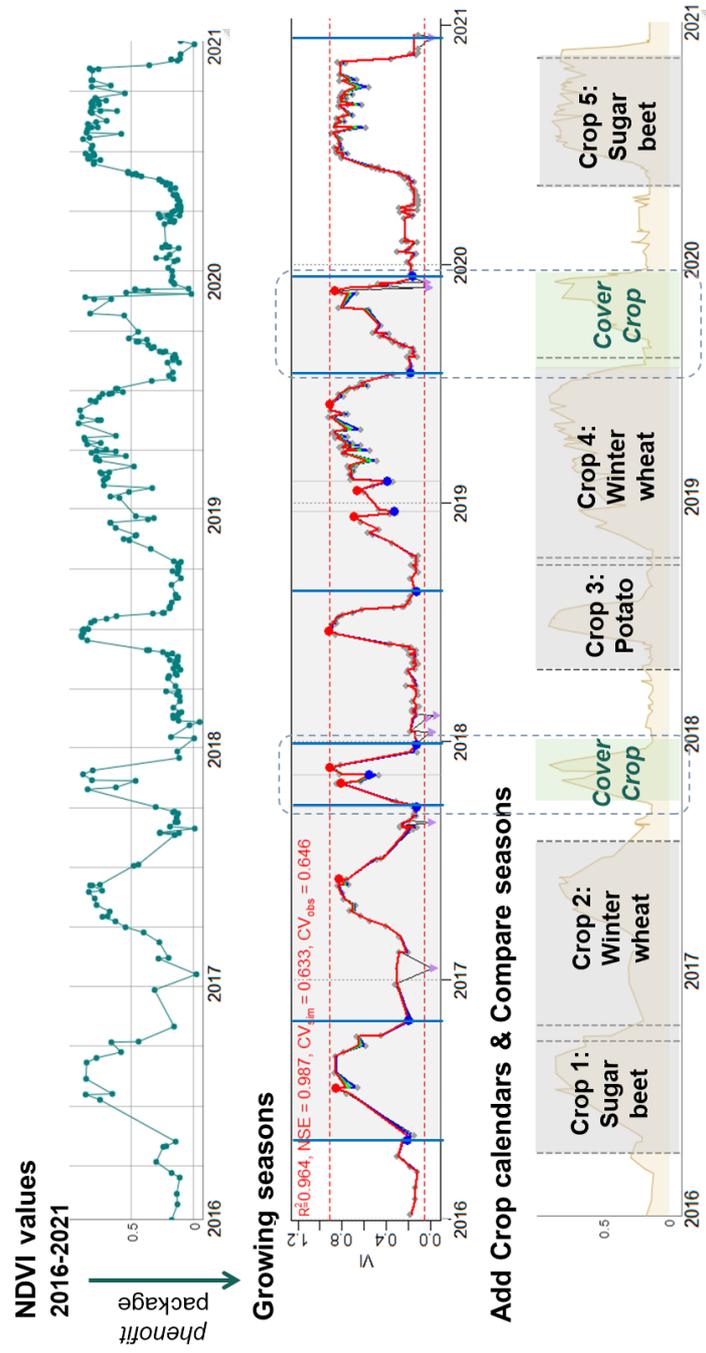


Figure A.3.8. Identifying cover crop of Lonzée station for period 2016-2021

Appendix 4 (Chapter 5)

Table A.4.1 SOC stock changes under Pedoclimate Regions (RSG: Reference Soil Group)

RSG in Climate zone (Atlantic)	Area(ha) /ratio	Clay (%)	SOC stock 2015 (t·ha ⁻¹)	SOC stock 2024 (t·ha ⁻¹)	Delta (t·ha ⁻¹)
Acrisols	25.02(0.05%)	13.26	40.43	41.64	1.21
Cambisols	3,411.84(6.74%)	15.76	51.72	51.33	-0.39
Fluvisols	1,057.18(2.09%)	13.068	46.13	45.93	-0.20
Luvvisols	46,161.49(91.13%)	13.33	46.17	46.37	0.19

Table A.4.2 SOC stock changes under soil association groups (Detail description of “Id” see Fig.5.3)

Id	Area (ha)	Clay (%)	SOC stock 2015 (t·ha ⁻¹)	SOC stock 2024 (t·ha ⁻¹)	Delta (t·ha ⁻¹)
31	15,279.85	13.05	44.84	44.95	0.11
35	19,762.52	13.40	46.39	46.94	0.56
38	1,410.35	12.80	44.62	44.73	0.12
33	4,516.03	14.56	51.12	51.05	-0.07
32	3,326.47	13.74	45.97	46.46	0.49
37	1,618.72	11.89	46.92	47.41	0.49
40	421.78	16.10	56.40	55.99	-0.41
60	530.00	13.71	47.24	47.49	0.24
30	1,496.31	15.61	44.02	44.53	0.51
34	921.93	14.67	51.30	51.08	-0.22
44	444.58	16.06	52.33	52.79	0.46
46	324.85	14.70	50.55	50.76	0.21
48	300.34	15.62	53.55	53.43	-0.13
41	184.45	14.66	51.24	50.21	-1.03
47	55.50	16.36	48.70	49.51	0.81
52	55.58	16.15	58.88	57.42	-1.46
42	6.27	11.71	48.31	46.92	-1.38

Bibliography

- Abatzoglou, J., Dobrowski, S., Parks, S., Hegewisch, K., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci Data* 5, 170191. <https://doi.org/10.1038/sdata.2017.191>
- Abiven, S., Menasseri, S. and Chenu, C. 2009. The effects of organic inputs over time on soil aggregate stability – A literature analysis. *Soil Biol. Biochem.* 41(1), pp.1-12.
- Adhikari, K., Kheir, R.B., Greve, M.B., Bøcher, P.K., Malone, B.P., Minasny, B., McBratney, A.B., Greve, M.H., 2013. High-resolution 3-D mapping of soil texture in Denmark. *Soil Science Society of America Journal* 77 (3), 860–876.
- Adhikari, K., Owens, P. R., Libohova, Z., Miller, D. M., Wills, S. A., Nemecek, J. 2019. Assessing soil organic carbon stock of Wisconsin, USA and its fate under future land use and climate change. *Sci. Total Environ.*, 667, 833-845
- Ahmed, Z., Shew, A., Nalley, L., Popp, M., Green, V.S., Brye, K., 2023. An examination of thematic research, development, and trends in remote sensing applied to conservation agriculture. *Soil Water Conserv. Res.* 12(1), 77-95.
- Aitkenhead, M.J., Coull, M.C., 2016. Mapping soil carbon stocks across Scotland using a neural network model. *Geoderma*, 262, 187-198.
- Ali, S., Begum, F., Hayat, R., Bohannan, B.J.M., 2017. Variation in soil organic carbon stock in different land uses and altitudes in Bagrot Valley, Northern Karakoram. *Acta Agric. Scand. B.* 67(6), 551–561.
- Allen, R., Tasumi, M., Trezza, R., 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—model. *J Irrig Drain Eng* 133(4):380–394.
- Alvarez-Vanhard, E., Corpetti, T., Houet, T. 2021. UAV satellite synergies for optical remote sensing applications: A literature review. *Science of Remote Sensing* 3, 100019. <https://doi.org/10.1016/j.srs.2021.100019>
- Alvear, M., Rosas, A., Rouanet, J., Borie, F. 2005. Effects of three soil tillage systems on some biological activities in an Ultisol from southern Chile. *Soil Tillage Res.* 82(2), 195–202. <https://doi.org/10.1016/j.still.2004.06.002>.
- Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Bol, R., Collins, C., Lal, R., Leifeld, J., Minasny, B., Pan, G., Paustian, K., Rumpel, C., Sanderman, J., van Groenigen, J. W., Mooney, S., van Wesemael, B., Wander, M., Chabbi, A. 2020. Towards a global-scale soil climate mitigation strategy. *Nat. Commun.*, 11, 5427.

- Amin, E., Verrelst, J., Rivera-Caicedo, J. P., Pipia, L., Ruiz-Verdú, A., Moreno, J., 2021. Prototyping Sentinel-2 green LAI and brown LAI products for cropland monitoring. *Remote Sens. Environ.* 255, 112168.
- Amundson, R., Berhe, A. A., Hopmans, J. W., Olson, C., Sztein, A. E., Sparks, D. L. 2015. Soil and human security in the 21st century. *Science*, 348(6235).
- Andries, A., Morse, S., Murphy, R.J., Lynch, J., Mota, B., Woolliams, E.R. 2021. Can Current Earth Observation Technologies Provide Useful Information on Soil Organic Carbon Stocks for Environmental Land Management Policy? *Sustainability*. 13(21), 12074. <https://doi.org/10.3390/su132112074>.
- Arrouays, D., Lagacherie, P., Hartemink, A., 2017. Digital soil mapping across the globe. *Geoderma Reg.* 9, 1–4. <https://doi.org/10.1016/j.geodrs.2017.03.002>.
- Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A., Morgan, C., Roudier, P., Poggio, L. and Mulder, V., 2020. Impressions of digital soil maps: The good, the not so good, and making them ever better. *Geoderma Regional*, 20, p.e 00255.
- Kempen, B., Dalsgaard, S., Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. *Eur. J. Soil Sci.* 47, 151–163. <https://doi.org/10.1111/j.1365-2389.1996.tb01386.x>.
- Ayuke, F.O., Kihara, J., Ayaga, G., Micheni, A.N., 2019. Conservation agriculture enhances soil fauna richness and abundance in low input systems: Examples from Kenya. *Frontiers in Environmental Science* 7.
- Azzari, G., Grassini, P., Edreira, J.I., Conley, S., Mourtzinis, S., Lobell, D.B., 2019. Satellite mapping of tillage practices in the north central US region from 2005 to 2016. *Remote Sens Environ.* 221, 417–429.
- Baghdadi, N., King, C., Bourguignon, A., Remond, A. 2002. Potential of ERS and Radarsat data for surface roughness monitoring over bare agricultural fields: Application to catchments in Northern France. *International Journal of Remote Sensing*, 23(17), 3427-3442.
- Baghdadi, N., El Hajj, M., Choker, M., Zribi, M., Bazzi, H., Vaudour, E., Gilliot, J.-M., Ebengo, D. M. 2018. Potential of Sentinel-1 images for estimating the soil roughness over bare agricultural soils. *Water*, 10(2), 131.
- Bai, X., Huang, Y., Ren, W., Coyne, M., Jacinthe, P.-A., Tao, B., Hui, D., Yang, J., Matocha, C. 2019. Responses of soil carbon sequestration to climate-smart agriculture practices: A meta-analysis. *Glob. Change Biol.*, 25(8), 2591-2606.
- Barbato, C.T., Strong, A.L. 2023. Farmer perspectives on carbon markets incentivizing agricultural soil carbon sequestration. *npj Clim. Action* 2, 26.
- Barnes, M.L., Yoder, L., Khodaei, M., 2021. Detecting winter cover crops and crop residues in the Midwest US using machine learning classification of thermal and optical imagery. *Remote Sens.* 13, 1998.

- Batjes, N. H., Ribeiro, E., van Oostrum, A., Leenaars, J., Hengl, T., Mendes de Jesus, J. (2017). Wosis: providing standardised soil profile data for the world. *Earth Syst. Sci. Data*, 9(1), 1–14. <https://doi.org/10.5194/essd-9-1-2017>
- Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. *Eur. J. Soil Sci.* 47, 151–163. <https://doi.org/10.1111/j.1365-2389.1996.tb01386.x>
- Batzias, F. A., Sidiras, D. K., Spyrou, E. K. 2005. Evaluating livestock manures for biogas production: a GIS based method. *Renew. Energy*, 30(8), 1161-1176.
- Bégué, A., Arvor, D., Bellon, B., Betbeder, J., De Aballeyra, D., Ferraz, R. P. D., Lebourgeois, V., Lelong, C., Simões, M., Verón, S. R. 2018. Remote sensing and cropping practices: A review. *Remote Sensing*, 10(1), 99.
- Beillouin, D., Corbeels, M., Demenois, J., Berre, D., Boyer, A., Fallot, A., Feder, F., Cardinael, R. 2023. A global meta-analysis of soil organic carbon in the Anthropocene. *Nature Communications*, 14, 3700.
- Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Soci, C., Villaume, S., Bidlot, J.-R., Haimberger, L., Woollen, J., Buontempo, C., Thépaut, J.-N. 2021. The ERA5 global reanalysis: Preliminary extension to 1950. *Q. J. R. Meteorol. Soc.*
- Bellamy, P., Loveland, P., Bradley, R. et al. Carbon losses from all soils across England and Wales 1978–2003. *Nature* 437, 245–248. 2005.
- Berihu, T., Girmay, G., Sebhatleab, M., Berhane, E., Zenebe, A. and Sigua, G., 2016. Soil carbon and nitrogen losses following deforestation in Ethiopia. *Agron Sustain Dev*, 37(1).
- Bhattacharya, S. S., Kim, K.-H., Das, S., Uchimiya, M., Jeon, B. H., Kwon, E., Szulejko, J. E. 2016. A review on the role of organic inputs in maintaining the soil carbon pool of the terrestrial ecosystem. *Journal of Environmental Management*, 167, 214-227. <https://doi.org/10.1016/j.jenvman.2015.09.042>
- Billings, S. A., Lajtha, K., Malhotra, A., Berhe, A. A., de Graaff, M.-A., Earl, S., Fraterrigo, J., Georgiou, K., Grandy, S., Hobbie, S. E., Moore, J. A. M., Nadelhoffer, K. N., et al. 2021. Soil organic carbon is not just for soil scientists: Measurement recommendations for diverse practitioners. *Ecological Applications*, 31(3), e02290. DOI: 10.1002/eap.2290.
- Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., Hostert, P., 2022. Mapping of crop types and crop sequences with combined time series of sentinel-1, sentinel-2 and Landsat 8 data for Germany. *Remote Sens Environ.* 269, 112831.
- Boegh, E., Soegaard, H., Broge, N., Hasager, C. B., Jensen, N. O., Schelde, K., Thomsen, A. 2002. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. *Remote Sens. Environ.*, 81(2–3), 179-193.

Bolinder, M.A., Crotty, F., Elsen, A., Frac, M., Kismányoky, T., Lipiec, J., Tits, M., Tóth, Z., Kätterer, T. 2020. The effect of crop residues, cover crops, manures and nitrogen fertilization on soil organic carbon changes in agroecosystems: a synthesis of reviews. *Mitig. Adapt. Strateg. Glob. Change* 25, 929–952.

Bolinder, M.A., Janzen, H.H., Gregorich, E.G., Angers, D.A., VandenBygaart, A.J. 2007. An approach for estimating net primary productivity and annual carbon inputs to soil for common agricultural crops in Canada. *Agric. Ecosyst. Environ.*, 118(1–4), 29–42. <https://doi.org/10.1016/j.agee.2006.05.013>

Bolinder, M.A., Kätterer, T., Poeplau, C., Börjesson, G., Parent, L.E., 2015. Net primary productivity and below-ground crop residue inputs for root crops: Potato (*Solanum tuberosum* L.) and sugar beet (*Beta vulgaris* L.). *Can. J. Soil Sci.* 95(2): 87–93. <https://doi.org/10.4141/cjss-2014-091>

Bousbih, S., Zribi, M., Pelletier, C., Gorrab, A., Lili-Chabaane, Z., Baghdadi, N., ben Aissa, N., Mougnot, B., 2019. Soil Texture Estimation Using Radar and Optical Data from Sentinel-1 and Sentinel-2. *Remote Sens.* 11(13), 1520.

Breiman, L., 2001. Random forests. *Mach Learn.* 45, 5–32.

Burger, R., Aouizerats, B., den Besten, N., Guillevic, P., Catarino, F., van der Horst, T., Jackson, D., Koopmans, R., Ridderikhoff, M., Robson, G., Zajdband, A., de Jeu, R. 2024. The Biomass Proxy: Unlocking Global Agricultural Monitoring through Fusion of Sentinel-1 and Sentinel-2. *Remote Sens.* 16(5), 835.

Buysse, P., Bodson, B., Debaq, A., De Ligne, A., Heinesch, B., Manise, T., Moureaux, C., Aubinet, M. 2017. Carbon budget measurement over 12 years at a crop production site in the silty-loam region in Belgium. *Agric. For. Meteorol.*, 246, 241–255.

Buysse, P., Roisin, C., Aubinet, M. 2013. Fifty years of contrasted residue management of an agricultural crop: Impacts on the soil carbon budget and on soil heterotrophic respiration. *Agric. Ecosyst. Environ.*, 167, 52–59. <https://doi.org/10.1016/j.agee.2013.01.006>

Cagnarini, C., Renella, G., Mayer, J., Hirte, J., Schulin, R., Costerousse, B., Della Marta, A., Orlandini, S., Menichetti, L. 2019. Multi-objective calibration of RothC using measured carbon stocks and auxiliary data of a long-term experiment in Switzerland. *Eur. J. Soil Sci.*, <https://doi.org/10.1111/ejss.12802>

Campbell, E. E., Paustian, K. 2015. Current developments in soil organic matter modeling and the expansion of model applications: A review. *Environ. Res. Lett.*, 10, 123004. <https://doi.org/10.1088/1748-9326/10/12/123004>.

Cárceles Rodríguez, B., Durán-Zuazo, V.H., Soriano Rodríguez, M., García-Tejero, I.F., Gálvez Ruiz, B., Cuadros Tavira, S., 2022. Conservation Agriculture as a sustainable system for Soil Health: A Review. *Soil Sys.* 6, 87.

- Castaldi, F., Chabrilat, S., Jones, A., Vreys, K., Bomans, B., van Wesemael, B., 2018. Soil Organic Carbon Estimation in Croplands by Hyperspectral Remote APEX Data Using the LUCAS Topsoil Database. *Remote Sens.* 10(2), 153. Error! Hyperlink reference not valid.
- Castaldi, F., Hueni, A., Chabrilat, S., Ward, K., Buttafuoco, G., Bomans, B., Vreys, K., Brell, M., van Wesemael, B., 2019. Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands. *ISPRS J. Photogramm. Remote Sens.* 147, 267–282.
- Castaldi, F., Buttafuoco, G., Bertinaria, F., Toscano, P. 2024. A geospatial approach for evaluating impact and potentiality of conservation farming for soil health improvement at regional and farm scale. *Soil and Tillage Research*, 244, 106212. <https://doi.org/10.1016/j.still.2024.106212>
- Castellanos-Navarrete, A., Rodríguez, A.C., de Goede, R.G.M., Kooistra, M.J., Sayre, K.D., Brussaard, L., Pulleman, M.M. Earthworm activity and soil structural changes under conservation agriculture in central Mexico. *Soil Tillage Res.* 2012, 123, 61–70.
- Chabbi, A., Lehmann, J., Ciais, P., Loescher, H. W., Cotrufo, M. F., Don, A., SanClements, M., Schipper, L., Six, J., Smith, P., Rumpel, C. 2017. Aligning agriculture and climate policy. *Nat. Clim. Chang.*, 7, 307–309.
- Chartin, C., Stevens, A., Goidts, E., Krüger, I., Carnol, M., van Wesemael, B., 2017. Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southern Belgium, Wallonia). *Geoderma Reg.* 9, 73–86. <https://doi.org/10.1016/j.geodrs.2016.12.006>.
- Chen, S., Arrouays, D., Leatitia Mulder, V., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., Richer-de-Forges, A. C., Walter, C., 2022. Digital mapping of GlobalSoilMap soil properties at a broad scale: A review. *Geoderma*, 409, 115567.
- Chen, S., Martin, M. P., Saby, N. P., Walter, C., Angers, D. A., Arrouays, D., 2018. Fine resolution map of top-and subsoil carbon sequestration potential in France. *Sci. Total Environ.* 630, 389-400.
- Chenu, C., Angers, D.A., Barré, P., Derrien, D., Arrouays, D., Balesdent, J., 2019. Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. *Soil Tillage Res.* 188, 41–52.
- Chong, J., 2022. Battle of the Ensemble—Random Forest vs Gradient Boosting. [online] Medium. Available at: <<https://towardsdatascience.com/battle-of-the-ensemble-random-forest-vs-gradient-boosting-6fbfed14cb7>>.
- Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., & Cox, P. M. 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geosci. Model Dev.*, 4, 701–722. <https://doi.org/10.5194/gmd-4-701-2011>

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S.V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* 219, 145–161.

Coleman, K. Jenkinson, D. S. 1996. RothC-26.3 - a model for the turnover of carbon in soil. Powlson, D. S., Smith, J. U. and Smith, P. (ed.) *Evaluation of soil organic matter models using existing long-term datasets.* Springer, Berlin. pp. 237-246

Coleman, K., Jenkinson, D.S., 2014. Rothc - A model for the turnover of carbon in soil - Model description and users guide. URL https://www.rothamsted.ac.uk/sites/default/files/RothC_guide_WIN.pdf (accessed 5.12.23).

Conchedda, G., Tubiello, F. N. 2020. Drainage of organic soils and GHG emissions: validation with country data. *Earth System Science Data*, 12, 3113-3131.

COWI, Ecologic Institute and IEEP. 2021. *Technical Guidance Handbook - setting up and implementing result-based carbon farming mechanisms in the EU Report to the European Commission, DG Climate Action, under Contract No. CLIMA/C.3/ETU/2018/007.* COWI, Kongens Lyngby.

Daughtry, C.S.T., Doraiswamy, P.C., Hunt, E.R., Stern, A.J., McMurtrey, J.E., Prueger, J.H., 2006. Remote sensing of crop residue cover and soil tillage intensity. *Soil Tillage Res.* 91, 101–108.

De Rosa, D., Ballabio, C., Lugato, E., Fasiolo, M., Jones, A., Panagos, P. 2023. Soil organic carbon stocks in European croplands and grasslands: How much have we lost in the past decade? *Glob. Change Biol.* <https://doi.org/10.1111/gcb.16992>

Dechow, R., Franko, U., Kätterer, T., Kolbe, H. 2019. Evaluation of the RothC model as a prognostic tool for the prediction of SOC trends in response to management practices on arable land. *Geoderma*, 337, 463-478.

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.

Derpsch, R., Kassam, A., Reicosky, D., Friedrich, T., Calegari, A., Basch, G., Gonzalez-Sanchez, E., Rheinheimer dos Santos, D. 2024. Nature's laws of declining soil productivity and Conservation Agriculture. *Soil Secur.*, 14, 100127.

Deutsch, C., and Journel, A., 1998. *GSLIB: Geostatistical Software Library and User's Guide.* New York: Oxford university press.

Djaman, K., Koudahe, K., Koubodana, H.D., Saibou, A., Essah, S., 2022. Tillage practices in potato (*solanum tuberosum* L.) production: A Review. *Am. J. Potato Res.* 99, 1–12.

Doblas-Rodrigo, Á., Gallejones, P., Artetxe, A., Merino, P. 2023. Role of livestock-derived amendments in soil organic carbon stocks in forage

- crops. *Sci. Total Environ.* 901, 165931 .
<https://doi.org/10.1016/j.scitotenv.2023.165931>
- Dong, T., Liu, J., Qian, B., He, L., Liu, J., Wang, R., Jing, Q., Champagne, C., McNairn, H., Powers, J., Shi, Y., Chen, J. M., Shang, J. 2020. Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data. *ISPRS J. Photogramm. Remote Sens.*, 168, 236-250.
- Dokuchaev, V.V., 1883. *Russian Chernozem.: Selected Works of V.V. Dokuchaev.* vol. I. Israel Program for Scientific Translations, Jerusalem (translated in 1967).
- 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.
- Dvorakova, K., Shi, P., Limbourg, Q., van Wesemael, B., 2020. Soil Organic Carbon Mapping from Remote Sensing: The Effect of Crop Residues. *Remote Sens.* 12(12), 1913.
- European Commission, 2021. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, EU Soil Strategy for 2030 – Reaping the benefits of healthy soils for people, food, nature and climate, COM (2021) 699 final. 17.11. Brussels, 17.11.2021.
- Evans, F., Shen, J., 2021. Long-Term Hindcasts of Wheat Yield in Fields Using Remotely Sensed Phenology, Climate Data and Machine Learning. *Remote Sens.* 13(13), 2435.
- Fantin, V., Buscaroli, A., Buttol, P., Novelli, E., Soldati, C., Zannoni, D., Zucchi, G., Righi, S. 2022. The RothC Model to Complement Life Cycle Analyses: A Case Study of an Italian Olive Grove. *Sustainability* 14(1), 569.
- FAO and ITPS. 2018. *Global Soil Organic Carbon Map (GSOCmap)* Technical Report. Rome. 162 pp.
- FAO, 2015. International Year of Soil. <http://www.fao.org/soils-2015/>.
- FAO, 2020. Technical specifications and country guidelines for Global Soil Organic Carbon Sequestration Potential Map (GSOCseq). Rome.
- FAO, IIASA. 2023. Harmonized World Soil Database version 2.0. Rome and Laxenburg. <https://doi.org/10.4060/cc3823en>
- FAO. 2020. GSOCseq Global Soil Organic Carbon Sequestration Potential Map Technical Manual. G. Peralta, L. Di Paolo, C. Omuto, K. Viatkin, I. Luotto, Y. Yigini, 1st Edition, Rome.
- FAO. 2011. *Save and Grow: A Policymaker’s Guide to the Sustainable Intensification of Smallholder Crop Production.* 17-19.
- Feigenwinter, I., Hörtnagl, L., Zeeman, M. J., Eugster, W., Fuchs, K., Merbold, L., & Buchmann, N. (2023). Large inter-annual variation in carbon sink strength of a permanent grassland over 16 years: Impacts of management practices and climate. *Agricultural and Forest Meteorology*, 340, 109613. <https://doi.org/10.1016/j.agrformet.2023.109613>

- Fendrich, A.N., Matthews, F., Van Eynde, E., Carozzi, M., Li, Z., d'Andrimont, R., Lugato, E., Martin, P., Ciais, P., Panagos, P., 2023. From regional to Parcel Scale: A high-resolution map of cover crops across Europe combining satellite data with statistical surveys. *Sci. Total Environ.* 873, 162300.
- Ferdinand M. S., Baret P. V. 2024. A method to account for diversity of practices in Conservation Agriculture – accepted. doi: 10.1007/s13593-024-00961-9. Manuscript number, ASDE-D-23-00503R1.
- Ferreira, C.S.S., Seifollahi-Aghmiuni, S., Destouni, G., Ghajarnia, N., Kalantari, Z., 2022. Soil degradation in the European Mediterranean region: Processes, status and consequences. *Sci. Total Environ.* 805, 150106.
- Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P., 2011. Solutions for a cultivated planet. *Nature* 478, 337–342.
- Follett, R. F., Stewart, C. E., Pruessner, E.G., Kimble, J.M., 2012. Effects of climate change on soil carbon and nitrogen storage in the US Great Plains. *J. Soil Water Conserv.* 67(5), 331–342. <https://doi.org/10.2489/jswc.67.5.331>.
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 29(5). <https://doi.org/10.1214/aos/1013203451>.
- Fujisaki, K., Chapuis-Lardy, L., Albrecht, A., Razafimbelo, T., Chotte, J. L., Chevallier, T., 2018. Data synthesis of carbon distribution in particle size fractions of tropical soils: Implications for soil carbon storage potential in croplands. *Geoderma*, 313, 41–51. <https://doi.org/10.1016/j.geoderma.2017.10.010>.
- Gabbrielli, M., Corti, M., Perfetto, M., Fassa, V., Bechini, L. 2022. Satellite-Based Frost Damage Detection in Support of Winter Cover Crops Management: A Case Study on White Mustard. *Agronomy* 12(9), 2025.
- Gale, M.R. Grigal, D.F. 1987. Vertical root distributions of northern tree species in relation to successional status. *Canadian Journal of Forest Research*, 17, 829–834.
- Gan, Y. T., Campbell, C. A., Janzen, H. H., Lemke, R. L., Basnyat, P., McDonald, C. L. 2009. Carbon input to soil from oilseed and pulse crops on the Canadian prairies. *Agric. Ecosyst. Environ.*, 132(3–4), 290–297.
- Garnier, J., Billen, G., Tournebise, J., Barré, P., Mary, B., Baudin, F. 2022. Storage or loss of soil active carbon in cropland soils: The effect of agricultural practices and hydrology. *Geoderma*, 407, 115538.
- Geremew, B., Tadesse, T., Bedadi, B., Gollany, H.T., Tesfaye, K., Aschalew, A., Tilaye, A., Abera, W. 2024. Evaluation of RothC model for predicting soil organic carbon stock in north-west Ethiopia. *Environ. Chall.* 15, 100909.

- Gholizadeh, A., Žižala, D., Saberioon, M., Borůvka, L., 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. *Remote Sens. Environ.* 218, 89–103.
- Giardina, C. P., Ryan, M. G., 2000. Evidence that decomposition rates of organic carbon in mineral soil do not vary with temperature. *Nature*. 404(6780), 858–861. <https://doi.org/10.1038/35009076>.
- Giardina, C.P., Ryan, M.G., Hubbard, R.M., Binkley, D., 2001. Tree Species and Soil Textural Controls on Carbon and Nitrogen Mineralization Rates. *Soil Sci. Soc. Am. J.* 65(4), 1272–1279. <https://doi.org/10.2136/sssaj2001.6541272x>.
- Giraldo-Perez, P., Raw, V., Greven, M., Goddard, M.R., 2021. A small effect of conservation agriculture on soil biodiversity that differs between biological kingdoms and geographic locations. *iScience* 24, 102280.
- Goidts, E. 2009. Soil organic carbon evolution at the regional scale: overcoming uncertainties & quantifying driving forces. Prom: van Wesemael, Bas <http://hdl.handle.net/2078.1/21726>
- Goffart, D., Curnel, Y., Planchon, V., Goffart, J.P., Defourny, P. (2021). Field-scale assessment of Belgian winter cover crops biomass based on Sentinel-2 data. *Eur. J. Agron.* 126, 126278. <https://doi.org/10.1016/j.eja.2021.126278>
- Goidts E, van Wesemael B. 2007. Regional assessment of the changes in soil organic carbon stocks of agricultural soils in southern Belgium between 1955 and 2005. *Geoderma* 141:341–354
- Goidts, E., Van Wesemael, B., Crucifix, M., 2009. Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales. *Eur. J. Soil Sci.* <https://doi.org/10.1111/j.1365-2389.2009.01157.x>
- Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*. 146(3–4), 403–411. <https://doi.org/10.1016/j.geoderma.2008.06.011>.
- Gong, Y.D, Xing, X.G., Wang, W.H., 2020. Factors determining soil water heterogeneity on the Chinese Loess Plateau as based on an empirical mode decomposition method. *J. Arid Land*. 12(3), 462–472.
- Goodman, J. M., Owens, P. R. 2012. Predicting soil organic carbon using mixed conceptual and geostatistical models. In B. Minasny, B. P. Malone, A. B. McBratney (Eds.), *Digital Soil Assessments and Beyond* (pp. 155-159). CRC Press, London.
- Gray, J. M., Wang, B., Waters, C. M., Orgill, S. E., Cowie, A. L., Ng, E. L. 2021. Digital mapping of soil carbon sequestration potential with enhanced vegetation cover over New South Wales, Australia. *Soil Use and Management*, 38(1), 587-598. DOI: 10.1111/sum.12766.
- Griffiths, P., Nendel, C., Hostert, P., 2019. Intra-annual reflectance composites from sentinel-2 and landsat for national-scale crop and land cover mapping. *Remote Sens Environ.* 220, 135–151.

- Gross, A., Glaser, B. 2021. Meta-analysis on how manure application changes soil organic carbon storage. *Scientific Reports*, 11, 5516.
- Guillaume, T., Bragazza, L., Levasseur, C., Libohova, Z., Sinaj, S. 2021. Long-term soil organic carbon dynamics in temperate cropland-grassland systems. *Agric. Ecosyst. Environ.*, 305, 107184.
<https://doi.org/10.1016/j.agee.2020.107184>
- Gutierrez, S., Grados, D., Møller, A. B., de Carvalho Gomes, L., Beucher, A. M., Giannini-Kurina, F., Wollesen de Jonge, L., Greve, M. H. 2023. Unleashing the sequestration potential of soil organic carbon under climate and land use change scenarios in Danish agroecosystems. *Sci. Total Environ.*, 905, 166921.
- Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene Selection for Cancer Classification using Support Vector Machines. *Mach Learn.* 46, 389–422.
- Haddaway, N.R., Hedlund, K., Jackson, L.E., Kätterer, T., Lugato, E., Thomsen, I.K., Jørgensen, H.B., Isberg, P.E., 2016. How does tillage intensity affect soil organic carbon? A systematic review protocol. *Environ. Evid.* 5(1).
- Helmick, J.L., Nauman, T.W., Thompson, J.A., 2014. Developing and assessing prediction intervals for soil property maps derived from legacy databases. 2014. In: Arrouays D., McKenzie N.J., Hempel J., Richer-de-Forges A.C., McBratney A.B. (eds), 2014. *GlobalSoilMap. Basis of the global soil information system*. Taylor Francis, CRC Press, London, p. 359–366.
- Hendriks, C., Lesschen, J.P., Timmermans, B., Hanegraaf, M., Dijkman, W., Rougoor, C., Cruijssen, J., Schepens, J., 2023. Description and development Soil Carbon Tool. Wageningen: Slim Landgebruik
- Heuvelink, G.B.M., Angelini, M.E., Poggio, L., Bai, Z., Batjes, N.H., van den Bosch, R., Bossio, D., Estella, S., Lehmann, J., et al. 2020. Machine learning in space and time for modelling soil organic carbon change. *Eur. J. Soil Sci.*, 71(6), 1010-1027. DOI: 10.1111/ejss.12998.
- Hijbeek R, van Loon MP, van Ittersum MK. 2019. Fertiliser use and soil carbon sequestration: opportunities and trade-offs. CCAFS Working Paper no. 264. Wageningen, the Netherlands: CGIARResearch Program on Climate Change, Agriculture and Food Security (CCAFS). Available online at: www.ccafs.cgiar.org
- Hively, W., Lamb, B., Daughtry, C., Shermeyer, J., McCarty, G., Quemada, M., 2018. Mapping crop residue and tillage intensity using worldview-3 satellite shortwave infrared residue indices. *Remote Sens.* 10, 1657.
- Hively, W. D., Lamb, B. T., Daughtry, C. S. T., Serbin, G., Dennison, P., Kokaly, R. F., Wu, Z., & Masek, J. G. 2021. Evaluation of SWIR crop residue bands for the Landsat Next mission. *Remote Sens.*, 13(18), 3718.

- Hobbs, P.R., Sayre, K., Gupta, R., 2007. The role of conservation agriculture in sustainable agriculture. *Philos. Trans. R. Soc. B: Biol. Sci.* 363(1491), 543–555.
- Hollis, J. M., Hannam, J., & Bellamy, P. H. (2011). Empirically-derived pedotransfer functions for predicting bulk density in European soils. *European Journal of Soil Science*.
- Holmstrom, D.A., Carter, M.R., 2000. Effect of subsoil tillage in the previous crop year on soil loosening and potato yield performance. *Can. J. Plant Sci.* 80, 161–164.
- Houghton, R., 2007. Balancing the global carbon budget. *Annu. Rev. Earth Planet. Sci.*, 35, 313-347.
<https://doi.org/10.1146/annurev.earth.35.031306.140057>.
- Hu, Q., Li, Z., Wang, L., Huang, Y., Wang, Y., Li, L., 2019. Rainfall Spatial Estimations: A Review from Spatial Interpolation to Multi-Source Data Merging. *Water* 11(3), 579. <https://doi.org/10.3390/w11030579>
- Hu, W., Si, B.C. 2013. Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition. *Geoderma*. 193–194, 180–188.
- Huang, H., Yang, L., Zhang, L., Pu, Y., Yang, C., Wu, Q., Cai, Y., Shen, F., Zhou, C. 2022. A review on digital mapping of soil carbon in cropland: progress, challenge, and prospect. *Environ. Res. Lett.*, 17(12), 123004.
- Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 454(1971), 903–995.
- Huang, X., Jia, Z., Jiao, X., Wang, J., Huang, X. 2022. Long-term manure applications to increase carbon sequestration and macroaggregate-stabilized carbon. *Soil Biol. Biochem.* 174, 108827.
<https://doi.org/10.1016/j.soilbio.2022.108827>
- INTERREG Carbon Farming project, 2021. Five promising measures to protect the climate by Carbon Farming. Available online at: <http://northsearegion.eu/media/16479/factsheets-top-5-carbon-farming-en.pdf>.
- IPCC, 2019. 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories, eds. E. Calvo Buendia, K. Tanabe, A. Kranjc, J. Baasansuren, M. Fukuda, S. Ngarize, A. Osako, Y. Pyrozhenko, P. Shermanau, S. Federici. IPCC, Switzerland.
- Jackson, R.B., Canadell, J., Ehleringer, J.R., Mooney, H.A., Sala, O.E. Schulze, E.D. 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia*, 108, 389–411.
- Jat, H.S., Datta, A., Choudhary, M., Sharma, P.C., Yadav, A.K., Choudhary, V., Gathala, M.K., Jat, M.L., McDonald, A., 2019. Climate smart agriculture practices improve soil organic carbon pools, biological

properties and crop productivity in cereal-based systems of north-west india. *CATENA* 181, 104059.

Jenny, H., 1941. *Factors of Soil Formation*. McGraw-Hill: New York, NY, USA.

Johnson, D.M., 2019. Using the Landsat Archive to map crop cover history across the United States. *Remote Sens Environ.* 232, 111286.

Johnson, R., Zhang, T., 2014. Learning Nonlinear Functions Using Regularized Greedy Forest. *IEEE Trans. Pattern Anal. Mach. Intell.* 36(5), 942–954.

Jordon, M.W., Smith, P., 2022. Modelling soil carbon stocks following reduced tillage intensity: A framework to estimate decomposition rate constant modifiers for RothC-26.3, demonstrated in north-west Europe. *Soil Tillage Res.* 222(105428).

Jordon, M.W., Smith, P., Long, P.R., Bürkner, P.C., Petrokofsky, G., Willis, K.J., 2022. Can Regenerative Agriculture increase national soil carbon stocks? Simulated country-scale adoption of reduced tillage, cover cropping, and ley-arable integration using RothC. *Sci. Total Environ.* 825(153955).

Kaonga, M.L., Coleman, K. 2008. Modelling soil organic carbon turnover in improved fallows in eastern Zambia using the RothC-26.3 model. *For. Ecol. Manage.* 256, 1160-1166.
<https://doi.org/10.1016/j.foreco.2008.06.017>

KC, K., Zhao, K., Romanko, M., Khanal, S., 2021. Assessment of the spatial and temporal patterns of cover crops using remote sensing. *Remote Sens.* 13, 2689.

Keel, S. G., Anken, T., Büchi, L., Chervet, A., Fliessbach, A., Flisch, R., Huguenin-Elie, O., Mäder, P., Mayer, J., Sinaj, S., Sturny, W., Wüst-Galley, C., Zihlmann, U., Leifeld, J. 2019. Loss of soil organic carbon in Swiss long-term agricultural experiments over a wide range of management practices. *Agric. Ecosyst. Environ.*, 286, 106654.
<https://doi.org/10.1016/j.agee.2019.106654>

Keel, S.G., Leifeld, J., Mayer, J., Taghizadeh-Toosi, A. and Olesen, J.E. 2017, Large uncertainty in soil carbon modelling related to method of calculation of plant carbon input in agricultural systems. *Eur J Soil Sci*, 68: 953-963.

Kempen, B., Dalsgaard, S., Kaaya, A., Chamuya, N., Ruipérez-González, M., Pekkarinen, A. and Walsh, M., 2019. Mapping topsoil organic carbon concentrations and stocks for Tanzania. *Geoderma*, 337, pp.164-180.

Keshavarz Afshar, R., Nilahyane, A., Chen, C., He, H., Bart Stevens, W., Iversen, W.M., 2019. Impact of conservation tillage and nitrogen on sugarbeet yield and quality. *Soil Tillage Res.* 191, 216–223.

Kharel, T.P., Bhandari, A.B., Mubvumba, P., Tyler, H.L., Fletcher, R.S., Reddy, K.N. 2023. Mixed-Species Cover Crop Biomass Estimation Using Planet Imagery. *Sensors*, 23(3), 1541. <https://doi.org/10.3390/s23031541>

- Kienast-Brown, S., Libohova, Z., Boettinger, J. 2022. Digital Soil Mapping. In *Soil Mapping and Process Modeling for Sustainable Land Use Management* (Chapter 5). USDA-NRCS & Utah State University.
- Kirschbaum, M. U. F. 1995. The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage. *Soil Biol. Biochem.*, 27(6), 753-760. [https://doi.org/10.1016/0038-0717\(94\)00242-S](https://doi.org/10.1016/0038-0717(94)00242-S)
- Kirschbaum, M.U. 2000. Will changes in soil organic carbon act as a positive or negative feedback on global warming? *Biogeochemistry* 48, 21–51.
- Kong, D., 2023. R phenofit: Extract Remote Sensing Vegetation Phenology Comprehensive. [R package phenofit version 0.3.8]. <https://cran.r-project.org/web/packages/phenofit/index.html>
- Kong, D., McVicar, T.R., Xiao, M., Zhang, Y., Peña-Arancibia, J.L., Filippa, G., Xie, Y., Gu, X., 2022. Phenofit : An R package for extracting vegetation phenology from time series remote sensing. *Methods Ecol. Evol.* 13, 1508–1527.
- Kontgis, C., Warren, M.S., Skillman, S.W., Chartrand, R., Moody, D.I., 2017. Leveraging sentinel-1 time-series data for mapping agricultural land cover and land use in the Tropics. 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp).
- 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>
- Krauss, M., Wiesmeier, M., Don, A., Cuperus, F., Gattinger, A., Gruber, S., Haagsma, W.K., Peigné, J., Chiodelli Palazzoli, M., Schulz, F., van der Heijden, M.G.A., Vincent-Caboud, L., Wittwer, R.A., Zikeli, S., Steffens, M. 2022. Reduced tillage in organic farming affects soil organic carbon stocks in temperate Europe. *Soil Tillage Res.*, 216, 105262. DOI: 10.1016/j.still.2021.105262.
- Kubitza, C., Krishna, V.V., Schulthess, U. et al. Estimating adoption and impacts of agricultural management practices in developing countries using satellite data. A scoping review. *Agron. Sustain. Dev.* 40, 16. 2020.
- Kuhn, M., Johnson, K., 2019. Feature selection overview. *Feature Engineering and Selection* 227–240.
- Kumar, P., Sajjad, H., Tripathy, B. R., Ahmed, R., Mandal, V. P., 2017. Prediction of spatial soil organic carbon distribution using Sentinel-2A and field inventory data in Sariska Tiger Reserve. *Nat. Hazards.* 90(2), 693–704. <https://doi.org/10.1007/s11069-017-3062-5>.
- Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote. Sens. Lett.* 14(5), 778–782.
- Kutsch, W.L., Aubinet, M., Buchmann, N., Smith, P., Osborne, B., Eugster, W., Wattenbach, M., Schruppf, M., Schulze, E.D., Tomelleri, E., Ceschia, E., Bernhofer, C., Béziat, P., Carrara, A., Di Tommasi, P.,

- Grünwald, T., Jones, M., Magliulo, V., Marloie, O., Moureaux, C., Ziegler, W., 2010. The net biome production of full crop rotations in Europe. *Agric. Ecosyst. Environ.* 139, 3(336-345).
<https://doi.org/10.1016/j.agee.2010.07.016>
- Laamrani, A., Joosse, P., McNairn, H., Berg, A., Hagerman, J., Powell, K., Berry, M., 2020. Assessing soil cover levels during the non-growing season using multitemporal satellite imagery and spectral unmixing techniques. *Remote Sens.* 12, 1397.
- Lal, R. 2004a. Soil carbon sequestration impacts on global climate change and food security. *Science*, 304(5677), 1623-1627.
- Lal, R. 2004b. Soil carbon sequestration to mitigate climate change. *Geoderma*, 123(1-2), 1-22. <https://doi.org/10.1016/j.geoderma.2004.01.032>
- Lal, R., 2015. Restoring Soil Quality to Mitigate Soil Degradation. *Sustainability*. 7, 5875-5895. <https://doi.org/10.3390/su7055875>.
- Lamichhane, S., Kumar, L., Wilson, B. 2019. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma*, 352, 395-413.
<https://doi.org/10.1016/j.geoderma.2019.05.031>
- Lawrence, C. R., Neff, J. C. Schimel, J. P. 2009. Does adding microbial mechanisms of decomposition improve soil organic matter models? A comparison of four models using data from a pulsed rewetting experiment. *Soil Biol. Biochem.* 41, 1923–1934.
- Le Noë, J., Manzoni, S., Abramoff, R., Bölscher, T., Bruni, E., Cardinael, R., Ciais, P., Chenu, C., Clivot, H., Derrien, D., Ferchaud, F., Garnier, P., Goll, D., Lashermes, G., Martin, M., Rasse, D., Rees, F., Sainte-Marie, J., Salmon, E., Schiedung, M., Schimel, J., Wieder, W., Abiven, S., Barré, P., Guenet, B. 2023. Soil organic carbon models need independent time-series validation for reliable prediction. *Commun. Earth Environ.*, 4, 158. <https://doi.org/10.1038/s43247-023-00830-5>
- Leenaars, JGB. 2013. Africa Soil Profiles Database, Version 1.1. A compilation of georeferenced and standardised legacy soil profile data for Sub-Saharan Africa (with dataset). Africa Soil Information Service (AfSIS) project. (ISRIC report; No. 2013/03). ISRIC. <https://edepot.wur.nl/261280>
- Lesschen, J. P., Hendriks, C., van de Linden, A., Timmermans, B., Keuskamp, J., Keuper, D., Hanegraaf, M., Conijn, S., Slier, T. 2020. Ontwikkeling praktijktool voor bodem C. Wageningen Environmental Research rapport; No. 2990. Wageningen Environmental Research. pp37-49. <https://doi.org/10.18174/517746>
- Li, B., Song, H., Cao, W., Wang, Y., Chen, J., Guo, J. 2021. Responses of soil organic carbon stock to animal manure application: A new global synthesis integrating the impacts of agricultural managements and environmental conditions. *Glob. Change Biol.*, 27(1), 90-104.
<https://doi.org/10.1111/gcb.15731>
- Li, S., Li, J., Li, C., Huang, S., Li, X., Li, S., Ma, Y. 2016. Testing the RothC and DNDC models against long-term dynamics of soil organic carbon

stock observed at cropping field soils in North China. *Soil Tillage Res.*, 163, 290–297.

Liang, Z.Z., Chen, S.C., Yang, Y.Y., Zhao, R.Y., Shi, Z., Viscarra Rossel, R.A., 2019. National digital soil map of organic matter in topsoil and its associated uncertainty in 1980's China. *Geoderma*. 335, 47–56.

Lin, C., Zhu, A.X., Wang, Z., Wang, X., Ma, R., 2020. The refined spatiotemporal representation of soil organic matter based on remote images fusion of Sentinel-2 and Sentinel-3. *Int. J. Appl. Earth Obs. Geoinf.* 89, 102094.

Liu, C., Zhang, Q., Tao, S., Qi, J., Ding, M., Guan, Q., Wu, B., Zhang, M., Nabil, M., Tian, F., Zeng, H., Zhang, N., Bavuudorj, G., Rukundo, E., Liu, W., Bofana, J., Beyene, A. N., Elnashar, A., 2020. A new framework to map fine resolution cropping intensity across the globe: Algorithm, validation, and implication. *Remote Sens. Environ.* 251, 112095.

Liu, F., Wu, H., Zhao, Y., Li, D., Yang, J.-L., Song, X., Shi, Z., Zhu, A.-X., Zhang, G.-L., 2022. Mapping high resolution National Soil Information Grids of China. *Sci. Bull.* 67(3), 328-340.
<https://doi.org/10.1016/j.scib.2021.10.013>.

Liu, Y., Wang, J., Yang, K., Ochir, A. 2024. Mapping livestock density distribution in the Selenge River Basin of Mongolia using random forest. *Sci. Rep.*, 14, Article 11090. <https://doi.org/10.1038/s41598-024-11090-3>

Loveland, P., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. *Soil Tillage Res.* 70(1), pp.1-18.

Lugato, E., Panagos, P., Bampa, F., Jones, A., Montanarella, L. 2013. A new baseline of organic carbon stock in European agricultural soils using a modelling approach. *Glob. Change Biol.*, 20(1), 313-324.
<https://doi.org/10.1111/gcb.12292>

Maas, E.D.v.L., Lal, R.A. 2023. A case study of the RothC soil carbon model with potential evapotranspiration and remote sensing model inputs. *Remote Sens. Appl. Soc. Environ.* 29, 100876.
<https://doi.org/10.1016/j.rsase.2022.100876>

Maillard, É., Angers, D. A. 2013. Animal manure application and soil organic carbon stocks: a meta-analysis. *Glob. Change Biol.*, 20(10), 666-679.

Malone, B., Hedley, C., Roudier, P., Minasny, B., Jones, E., McBratney, A. B. 2018. Auditing on-farm soil carbon stocks using downscaled national mapping products: Examples from Australia and New Zealand. *Geoderma Reg.* 13, 1–14.

Malone, B., McBratney, A., Minasny, B., 2011. Empirical estimates of uncertainty for mapping continuous depth functions of soil attributes. *Geoderma*, 160(3–4), 614–626.
<https://doi.org/10.1016/j.geoderma.2010.11.013>

- Malone, B., Minasny, B., McBratney, A.B., 2016. Using R for Digital Soil Mapping (Progress in Soil Science) (1st ed. 2017 ed.). Springer. P178-187.
- Malone, B., Styc, Q., Minasny, B., McBratney, A.B., 2017. Digital soil mapping of soil carbon at the farm scale: a spatial downscaling approach in consideration of measured and uncertain data. *Geoderma*. 290, 91-99.
- Manrique, L.A., Jones, C.A. 1991. Bulk density of soils in relation to soil physical and chemical properties. *Soil Sci Soc Am J* 55, 476–481.
- Manzoni, S., Porporato, A. 2009. Soil carbon and nitrogen mineralization: Theory and models across scales. *Soil Biol. Biochem.*, 41(7), 1355–1379.
- Martin, M.P., Dimassi, B., Román Dobarco, M., Guenet, B., Arrouays, D., Angers, D. A., Blache, F., Huard, F., Soussana, J., Pellerin, S., 2021. Feasibility of the 4 per 1000 aspirational target for soil carbon: A case study for France. *Glob. Change Biol.* 27(11), 2458–2477. <https://doi.org/10.1111/gcb.15547>.
- Mazzoncini, M., Sapkota, T. B., Bàrberi, P., Antichi, D., Risaliti, R. 2011. Long-term effect of tillage, nitrogen fertilization and cover crops on soil organic carbon and total nitrogen content. *Soil Tillage Res.*, 114(2), 165-174.
- McBratney, A.B., Mendonca Santos, M.L., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4).
- Meersmans, J., van Wesemael, B., Goidts, E., van Molle, M., De Baets, S., De Ridder, F., 2011. Spatial analysis of soil organic carbon evolution in Belgian croplands and grasslands, 1960–2006. *Glob. Chang. Biol.* 17, 466–479.
- Meyer, H., Reudenbach, C., Hengl, T., Katurji, M. and Naus, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environ. Model. Softw.* 101, pp.1-9.
- Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B. S., Field, D. J., Gimona, A., Hedley, C. B., Hong, S. Y., Mandal, B., Marchant, B. P., Martin, M., McConkey, B. G., Mulder, V. L., O'Rourke, S., Winowiecki, L. 2017. Soil carbon 4 per mille. *Geoderma*, 292, 59-86. DOI: 10.1016/j.geoderma.2017.01.002.
- Minasny, B., McBratney, A. B. 2016. Digital soil mapping: A brief history and some lessons. *Geoderma*, 264(Part B), 301-311.
- Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput. Geosci.* 32(9): 1378-1388. <https://doi: 10.1016/j.cageo.2005.12.009>.
- Mishra, U., Gautam, S., Riley, W.J., Hoffman, F.M., 2020. Ensemble Machine Learning Approach Improves Predicted Spatial Variation of

Surface Soil Organic Carbon Stocks in Data-Limited Northern Circumpolar Region. *Frontiers in Big Data*. 3. <https://doi.org/10.3389/fdata.2020.528441>.

Morais, T.G., Teixeira, R.F., Domingos, T. 2019. Detailed global modelling of soil organic carbon in cropland, grassland and forest soils. *PLoS ONE*, 14(9), e0222604. <https://doi.org/10.1371/journal.pone.0222604>

Morvan, X., Saby, N.P.A., Arrouays, D., Le Bas, C., Jones, R.J.A., Verheijen, F.G.A., Bellamy, P.H., Stephens, M., Kibblewhite, M.G. 2008. Soil monitoring in Europe: A review of existing systems and requirements for harmonisation. *Sci. Total Environ.*, 391(1), 1-12.

Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Arrouays, D., 2016. GlobalSoilMap France: high-resolution spatial modelling the soils of France up to two meter depth. *Sci Total Environ*. 573:1352–1369.

Mzid, N., Castaldi, F., Tolomio, M., Pascucci, S., Casa, R., Pignatti, S., 2022. Evaluation of agricultural bare soil properties retrieval from Landsat 8, sentinel-2 and Prisma Satellite Data. *Remote Sens*. 14, 714.

Naab, J.B., Mahama, G.Y., Yahaya, I., Prasad, P.V., 2017. Conservation agriculture improves soil quality, crop yield, and incomes of smallholder farmers in North Western Ghana. *Frontiers in Plant Science* 8.

Nadeu, E., Gobin, A., Fiener, P., van Wesemael, B., van Oost, K. 2015. Modelling the impact of agricultural management on soil carbon stocks at the regional scale: the role of lateral fluxes. *Global Change Biology*, 21(6), 2132-2147.

Nagler, P.L., Inoue, Y., Glenn, E.P., Russ, A.L., Daughtry, C.S.T., 2003. Cellulose absorption index (CAI) to quantify mixed soil–plant litter scenes. *Remote Sens Environ*. 87, 310–325.

Natekin, A., Knoll, A., 2013. Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7. <https://doi.org/10.3389/fnbot.2013.00021>.

Nguyen, M.D., Baez-Villanueva, O.M., Bui, D.D., Nguyen, P.T., Ribbe, L., 2020. Harmonization of Landsat and Sentinel 2 for Crop Monitoring in Drought Prone Areas: Case Studies of Ninh Thuan (Vietnam) and Bekaa (Lebanon). *Remote Sens*. 12(2), 281. <https://doi.org/10.3390/rs12020281>.

Ocholla, I. A., Pellikka, P., Karanja, F., Vuorinne, I., Väisänen, T., Boitt, M., Heiskanen, J. 2024. Livestock detection and counting in Kenyan rangelands using aerial imagery and deep learning techniques. *Remote Sens.*, 16(16), 2929. <https://doi.org/10.3390/rs16162929>

Osborne, B., Saunders, M., Walmsley, D., Jones, M., Smith, P. 2010. Key questions and uncertainties associated with the assessment of the cropland greenhouse gas balance. *Agriculture, Ecosystems & Environment*, 139(3), 293-301. <https://doi.org/10.1016/j.agee.2010.05.009>

Osinaga, N., Álvarez, C. and Taboada, M., 2018. Effect of deforestation and subsequent land use management on soil carbon stocks in the South American Chaco. *SOIL*, 4(4), pp.251-257.

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

- Padarian, J., Minasny, B., McBratney, A.B., 2017. Chile and the Chilean soil grid: a contribution to GlobalSoilMap. *Geoderma Reg.* 9, 17–28.
- Page, K.L., Dang, Y.P., Dalal, R.C. 2020. The ability of conservation agriculture to conserve soil organic carbon and the subsequent impact on soil physical, chemical, and biological properties and yield. *Front. sustain. food syst.* 4.
- Palm, C., Blanco-Canqui, H., DeClerck, F., Gatere, L., Grace, P., 2014. Conservation agriculture and ecosystem services: An overview. *Agric. Ecosyst. Environ.* 187, 87–105. <https://doi.org/10.1016/j.agee.2013.10.010>.
- Pan, G., Xu, X., Smith, P., Pan, W., Lal, R. 2010. An increase in topsoil SOC stock of China's croplands between 1985 and 2006 revealed by soil monitoring. *Agric. Ecosyst. Environ.*, 136(1–2), 133–138. DOI: 10.1016/j.agee.2009.12.011.
- Panagos, P., Standardi, G., Borrelli, P., Lugato, E., Montanarella, L., Bosello, F., 2018. Cost of agricultural productivity loss due to soil erosion in the European Union: From direct cost evaluation approaches to the use of macroeconomic models. *Land Degrad Dev.* 29, 471–484.
- Pastorello, G., Trotta, C., Canfora, E. et al., 2018. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Sci. Data* 7 (225). <https://doi.org/10.1038/s41597-020-0534-3>
- Paustian, K., Andr n, O., Janzen, H.H., Lal, R., Smith, P., Tian, G., Tiessen, H., Noordwijk, M., Woomer, P.L., 1997. Agricultural soils as a sink to mitigate CO₂ emissions. *Soil Use Manag.* 13(s4), 230–244.
- Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G. P., Smith, P. 2016. Climate-smart soils. *Nature*, 532, 49–57. <https://doi.org/10.1038/nature17174>
- Peigné, J., Ball, B.C., Roger-Estrade, J., David, C., 2007. Is conservation tillage suitable for organic farming? A Review. *Soil Use Manag.* 23, 129–144.
- Pellerin, S., Bami re, L., Angers, D., B line, F., Benoit, M., Butault, J.-P., Chenu, C., Colnenne-David, C., De Cara, S., Delame, N., Doreau, M., Dupraz, P., Faverdin, P., Garcia-Launay, F., Hassouna, M., H nault, C., Jeuffroy, M.-H., Klumpp, K., Metay, A., Moran, D., Recous, S., Samson, E., Savini, I., Pardon, L., Chemineau, P. 2017. Identifying cost-competitive greenhouse gas mitigation potential of French agriculture. *Environmental Science Policy*, 77, 130–139. DOI: 10.1016/j.envsci.2017.08.003.
- Pesce, S., Balugani, E., De Paz, J. M., Marazza, D., Visconti, F. 2024. A Modified Version of RothC to Model the Direct and Indirect Effects of Rice Straw Mulching on Soil Carbon Dynamics, Calibrated in Two Valencian Citrus Orchards. *Soil Syst.*, 8(1), 12. <https://doi.org/10.3390/soilsystems8010012>
- Petito, M., Cantalamessa, S., Pagnani, G., Degiorgio, F., Parisse, B., Pisante, M., 2022. Impact of conservation agriculture on soil erosion in the annual cropland of the Apulia region (southern Italy) based on the Rusle-GIS-Gee Framework. 12, 281.

Phelan, L., Chapman, P. J., Ziv, G. 2024. The emerging global agricultural soil carbon market: The case for reconciling farmers' expectations with the demands of the market. *Environmental Development*, 49, 100941. DOI: 10.1016/j.envdev.2023.100941.

Piedelobo, L., Hernández-López, D., Ballesteros, R., Chakhar, A., del Pozo, S., González-Aguilera, D., Moreno, M.A., 2019. Scalable pixel-based crop classification combining Sentinel-2 and Landsat-8 data time series: Case study of the Duero river basin. *Agric. Syst.* 171, 36–50.

Pique, G., Fieuzal, R., Al Bitar, A., Veloso, A., Tallec, T., Brut, A., Ferlicoq, M., Zawilski, B., Dejoux, J.-F., Gibrin, H., Ceschia, E. 2020. Estimation of daily CO₂ fluxes and of the components of the carbon budget for winter wheat by the assimilation of Sentinel 2-like remote sensing data into a crop model. *Geoderma*, 376, 114428.

Pittman, S.J., Brown, K.A., 2011. Multi-Scale Approach for Predicting Fish Species Distributions across Coral Reef Seascapes. *PLoS ONE*. 6(5), e20583.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S. and Péliissier, R., 2020. Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nat. Commun.* 11(1).

Poepplau, C., Jacobs, A., Don, A., Vos, C., Schneider, F., Wittnebel, M., Tiemeyer, B., Heidkamp, A., Prietz, R., Flessa, H. 2020. Stocks of organic carbon in German agricultural soils—Key results of the first comprehensive inventory. *J. Plant Nutr. Soil Sci.* 183(6), 665–676.

Poepplau, C., Dechow, R. 2023. The legacy of one hundred years of climate change for organic carbon stocks in global agricultural topsoils. *Sci. Rep.*, 13, 7483.

Poepplau, C., Don, A. 2013. Sensitivity of soil organic carbon stocks and fractions to different land-use changes across Europe. *Geoderma*, 192, 189–201.

Poepplau, C., Don, A. 2015. Carbon sequestration in agricultural soils via cultivation of cover crops – A meta-analysis. *Agric. Ecosyst. Environ.*, 200, 33–41.

Poggio, L., de Sousa, L., Batjes, N., Heuvelink, G., Kempen, B., Ribeiro, E. and Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, 7(1), pp.217–240.

Porwollik, V., Rolinski, S., Heinke, J., von Bloh, W., Schaphoff, S., Müller, C. 2022. The role of cover crops for cropland soil carbon, nitrogen leaching, and agricultural yields – a global simulation study with LPJmL (V. 5.0-tillage-cc). *Biogeosciences*, 19, 957–973. <https://doi.org/10.5194/bg-19-957-2022>

Pouladi, N., Gholizadeh, A., Khosravi, V., Borůvka, L. 2023. Digital mapping of soil organic carbon using remote sensing data: A systematic

review. *CATENA*, 232, 107409.

<https://doi.org/10.1016/j.catena.2023.107409>

Powlson, D. S., Whitmore, A. P., Goulding, K. W. T. 2011. Soil carbon sequestration to mitigate climate change: a critical re-examination to identify the true and the false. *Soil Use Manage.*, 27(2), 107-117.

Pradhan, A., Chan, C., Roul, P. K., Halbrecht, J., Sipes, B. 2018. Potential of conservation agriculture (CA) for climate change adaptation and food security under rainfed uplands of India: a transdisciplinary approach. *Agric. Syst.*

Priori, S., Zanini, M., Falcioni, V., Casa, R. 2024. Topsoil vertical gradient in different tillage systems: An analytical review. *Soil Tillage Res.*, 236, 105947.

Quemada, M., Hively, W.D., Daughtry, C.S.T., Lamb, B.T., Shermeyer, J., 2018. Improved crop residue cover estimates obtained by coupling spectral indices for residue and moisture. *Remote Sens Environ.* 206, 33–44.

Raffeld, A. M., Bradford, M. A., Jackson, R. D., Rath, D., Sanford, G. R., Tautges, N., Oldfield, E. E. 2024. The importance of accounting method and sampling depth to estimate changes in soil carbon stocks. *Carbon Balance Manage.*, 19, Article 2. <https://doi.org/10.1186/s13021-024-0023-9>

Rehberger, E., West, P.C., Spillane, C., McKeown, P.C., 2023. What climate and environmental benefits of regenerative agriculture practices? an evidence review. *Environ. res. commun.* 5, 052001.

Robinson, D.A., Emmett, B.A., Reynolds, B., Rowe, E.C., Spurgeon, D., Keith, A.M., et al., 2012. Soil natural capital and ecosystem service delivery in a world of global soil change. In: Hester, R.E., Harrison, R.M. (Eds.), *Soils and Food Security*. Royal Society of Chemistry, Cambridge, pp. 41–68.

Romanenkov, V., Belichenko, M., Petrova, A., Raskatova, T., Jahn, G., & Krasilnikov, P. 2019. Soil organic carbon dynamics in long-term experiments with mineral and organic fertilizers in Russia. *Geoderma Reg.*, 17, e00221.

Rosinger, C., Keiblinger, K., Bieber, M., Bernardini, L.G., Huber, S., Mentler, A., Sae-Tun, O., Scharf, B., Bodner, G. 2023. On-farm soil organic carbon sequestration potentials are dominated by site effects, not by management practices. *Geoderma*, 116466. DOI: [10.1016/j.geoderma.2023.116466](https://doi.org/10.1016/j.geoderma.2023.116466)

Roudier, P., Hewitt, A., Beaudette, D., 2012. A conditioned Latin hypercube sampling algorithm incorporating operational constraints. *Digital Soil Assessments and Beyond*. 227–231. <https://doi.org/10.1201/b12728-46>.

Rumpel, C., Amiraslani, F., Bossio, D., Chenu, C., Garcia Cardenas, M., Henry, B., Fuentes Espinoza, A., Koutika, L.-S., Ladha, J., Madari, B. E., Minasny, B., Olaleye, A. (O.), Sall, S. N., Shirato, Y., Soussana, J.-F., Varela-Ortega, C. 2023. Studies from global regions indicate promising avenues for maintaining and increasing soil organic carbon stocks. *Reg Environ Change* 23, 8 2023.

- Rumpel, C., Amiraslani, F., Chenu, C., Garcia Cardenas, M., Kaonga, M., Koutika, L.-S., Ladha, J., Madari, B., Shirato, Y., Smith, P., Soudi, B., Soussana, J.-F., Whitehead, D., & Wollenberg, E. 2020. The 4p1000 initiative: Opportunities, limitations and challenges for implementing soil organic carbon sequestration as a sustainable development strategy. *Ambio*, 49(1), 350–360.
- Saby, N.P.A., Bellamy, P.H., Morvan, X., Arrouays, D., Jones, R.J.A., Verheijen, F.G.A., Kibblewhite, M.G., Verdoodt, A., Berényi Üveges, J., Freudenschuß, A., Simota, C. 2008a. Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? *Glob. Change Biol.*, 14, 2432-2442.
- Saby, N. P. A., Arrouays, D., Antoni, V., Lemercier, B., Follain, S., Walter, C., Schwartz, C. 2008b. Changes in soil organic carbon in a mountainous French region, 1990–2004. *Soil Use and Management*, 24(3), 254–262.
- Sanderman J, Hengl T, Fiske GJ. 2017. Soil carbon debt of 12,000 years of human land use. *Proceedings of the National Academy of Sciences of the United States of America.*;114(36):9575–9580. doi: 10.1073/pnas.1706103114.
- Satalino, G., Mattia, F., Balenzano, A., Lovergine, F.P., Rinaldi, M., de Santis, A.P., Ruggieri, S., Nafria Garcia, D.A., Gomez, V.P., Ceschia, E., Planells, M., Toan, T.L., Ruiz, A., Moreno, J. 2018. Sentinel-1 & Sentinel-2 Data for Soil Tillage Change Detection. *IGARSS 2018 - 2018 IEEE Int. Geosci. Remote Sens. Symp.*
- Scarlat, N., Fahl, F., Dallemand, J.-F., Monforti, F., Motola, V., 2018. A spatial analysis of biogas potential from manure in Europe. *Renew. Sustain. Energy Rev.* 94 (915-930). <https://doi.org/10.1016/j.rser.2018.06.035>
- Schaefer, M. V., Bogie, N. A., Rath, D., Marklein, A. R., Garniwan, A., Haensel, T., Lin, Y., Avila, C. C., Nico, P. S., Scow, K. M., Brodie, E. L., Riley, W. J., Fogel, M. L., Berhe, A. A., Ghezzehei, T. A., Parikh, S., Keiluweit, M., Ying, S. C. 2020. Effect of Cover Crop on Carbon Distribution in Size and Density Separated Soil Aggregates. *Soil Syst.*, 4(1), 6. <https://doi.org/10.3390/soilsystems4010006>
- Scharlemann, J. P., Tanner, E. V., Hiederer, R., Kapos, V. 2014. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, 5(1), 81–91. <https://doi.org/10.4155/cmt.13.77>
- Seitz, D., Fischer, L.M., Dechow, R., Wiesmeier, M., Don, A., 2023. The potential of cover crops to increase soil organic carbon storage in German croplands. *Plant Soil* 488, 157-173. <https://doi.org/10.1007/s11104-022-05438-w>
- Selin Noren, I., van Geel, W., de Haan, J. 2021. Cover crop reference values: effective organic matter and nitrogen uptake, Wageningen Research, Report WPR 877.
- Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A., Randall, D.A., 1996. A revised land surface parameterization (SiB2)

- for atmospheric GCMS. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Clim.*, 9, 706–737.
[https://doi.org/10.1175/1520-0442\(1996\)009<0706:ARLSPF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<0706:ARLSPF>2.0.CO;2).
- Senapati, N., Hulugalle, N.R., Smith, P., Wilson, B.R., Yeluripati, J.B., Daniel, H., Ghosh, S., Lockwood, P. 2014. Modelling soil organic carbon storage with RothC in irrigated Vertisols under cotton cropping systems in the sub-tropics. *Soil Tillage Res.* 143, 38-49.
<https://doi.org/10.1016/j.still.2014.05.009>
- Sharma, B., 2015. What are the advantages/disadvantages of using Gradient Boosting over Random Forests?. [online] Quora. Available at: <<https://qr.ae/pv2WX2>>.
- Sharma, S., Dhakal, K., Wagle, P., Kilic, A., 2020. Retrospective tillage differentiation using the landsat-5 TM archive with discriminant analysis. *Agrosyst. geosci. environ.* 3.
- Shen, Y., Xiao, Z., Wang, Y., Yao, L., Xiao, W. 2022). Multisource remote sensing based estimation of soil NO_x emissions from fertilized cropland at high-resolution: spatio-temporal patterns and impacts. *J. Geophys. Res. Atmos.*, 127(19), e2022JD036741.
<https://doi.org/10.1029/2022JD036741>
- Sherrod, L.A., Dunn, G., Peterson, G.A., Kolberg, R.L., 2002. Inorganic Carbon Analysis by Modified Pressure-Calcimeter Method. *Soil Sci. Soc. Am. J.* 66(1), 299.
- Shi, P., Castaldi, F., van Wesemael, B. and Van Oost, K., 2020. Vis-NIR spectroscopic assessment of soil aggregate stability and aggregate size distribution in the Belgian Loam Belt. *Geoderma*, 357, p.113958.
- Shirato, Y., Taniyama, I. 2003. Testing the suitability of the Rothamsted Carbon model for long-term experiments on Japanese non-volcanic upland soils. *Soil Science and Plant Nutrition* 49(6), 921-925.
- Sierra, C. A., Müller, M., and Trumbore, S. E. 2014. Modeling radiocarbon dynamics in soils: SoilR version 1.1, *Geosci. Model Dev.* 7, 1919–1931.
- Sithole, N.J., Magwaza, L.S., Thibaud, G.R., 2019. Long-term impact of no-till conservation agriculture and N-fertilizer on soil aggregate stability, infiltration and distribution of C in different size fractions. *Soil Tillage Res.* 190, 147–156.
- Smith, J., Smith, P., Wattenbach, M., Zaehle, S., Hiederer, R., Jones, R. J. A., Montanarella, L., Rounsevell, M. D. A., Reginster, I., Ewert, F. 2005. Projected changes in mineral soil carbon of European croplands and grasslands, 1990–2080. *Glob. Change Biol.*, 11(12), 2141-2152.
- Smith, P., Fang, C., Dawson, J.J.C., Moncrieff, J.B., 2008. Impact of Global Warming on Soil Organic Carbon. *Adv. Agron.* 97, 1–43.
- Smith, P., Lanigan, G., Kutsch, W.L., Buchmann, N., Eugster, W., Aubinet, M., Ceschia, E., Béziat, P., Yeluripati, J.B., Osborne, B., Moors, E.J., Brut, A., Wattenbach, M., Saunders, M., Jones, M. 2010. Measurements

necessary for assessing the net ecosystem carbon budget of croplands. *Agric. Ecosyst. Environ.* 139, 302–315. <https://doi.org/10.1016/j.agee.2010.04.004>

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., Scholes, B., Sirotenko, O., 2007. Chapter 8: Agriculture. in: Metz, B., Davidson, O.R., Bosch, P.R., Dave, R., Meyer L.A. (Eds.), *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., McCarl, B., Ogle, S., O'Mara, F., Rice, C., Scholes, B., Sirotenko, O., Howden, M., McAllister, T., Pan, G., Romanenkov, V., Schneider, U., Towprayoon, S., Wattenbach, M., Smith, J. 2008. Greenhouse gas mitigation in agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1492), 789-813.

Smith, P., Soussana, J.-F., Angers, D., Schipper, L., Chenu, C., Rasse, D.P., Batjes, N.H., van Egmond, F., McNeill, S., Kuhnert, M., Arias-Navarro, C., Olesen, J.E., Chirinda, N., Fornara, D., Wollenberg, E., Álvaro-Fuentes, J., Sanz-Cobena, A., Klumpp, K. 2019. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Glob. Change Biol.* 25, 1906–1924. <https://doi.org/10.1111/gcb.14815>

Soil Service of Belgium, Ghent University, 2006. Develop an expert system for advising carbon management in agricultural soils. (Originally published in Dutch: Ontwikkelen van een expertsysteem voor het adviseren van het koolstofbeheer in de landbouwbodems). <https://publicaties.vlaanderen.be/view-file/20207>.

Sonmez, N.K., Slater, B., 2016. Measuring intensity of tillage and plant residue cover using remote sensing. *Eur. J. Remote Sens.* 49, 121–135.

Sparks, A.M., Bouhamed, I., Boschetti, L., Gitas, I.Z., Kalaitzidis, C., 2022. Mapping arable land and permanent agriculture extent and change in southern Greece using the European Union Lucas Survey and a 35-year Landsat Time Series analysis. *Remote Sens.* 14, 3369.

Stallone, A., Cicone, A., Materassi, M., 2020. New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms. *Sci. Rep.* 10(1). <https://doi.org/10.1038/s41598-020-72193-2>.

Stevens, F., Bogaert, P., van Wesemael, B., 2015. Spatial filtering of a legacy dataset to characterize relationships between soil organic carbon and soil texture. *Geoderma.* 237-238, 224–236. <https://doi.org/10.1016/j.geoderma.2014.09.005>.

Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., de Remy de Courcelles, V., Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., Lehmann,

- O'Donnell, J., Parton, W., Whitehead, D., Zimmermann, M. 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems Environment*, 164, 80-99. DOI: 10.1016/j.agee.2012.10.001.
- Taghizadeh-Toosi, A., Christensen, B.T., Glendining, M., Olesen, J.E. 2016. Consolidating soil carbon turnover models by improved estimates of belowground carbon input. *Sci. Rep.*, 6, 32568. <https://doi.org/10.1038/srep32568>
- Taghizadeh-Toosi, A., Olesen, J.E., Kristensen, K., Elsgaard, L., Østergaard, H.S., Lægdsmand, M., Greve, M.H., Christensen, B.T. 2014. Changes in carbon stocks of Danish agricultural mineral soils during 1986–2009: Effects of management. *Eur. J. Soil Sci.*, 65, 730-740.
- Taghizadeh-Toosi, A., Olesen, J. E. 2016. Modelling soil organic carbon in Danish agricultural soils suggests low potential for future carbon sequestration. *Agric. Syst.*, 145, 83-89. <https://doi.org/10.1016/j.agry.2016.03.004>
- Tashi, S., Singh, B., Keitel, C., Adams, M., 2016. Soil carbon and nitrogen stocks in forests along an altitudinal gradient in the eastern Himalayas and a meta-analysis of global data. *Glob. Change Biol.* 22(6), 2255–2268.
- Thieme, A., Yadav, S., Oddo, P.C., Fitz, J.M., McCartney, S., King, L etc, 2020. Using NASA Earth observations and Google Earth Engine to map winter cover crop conservation performance in the Chesapeake Bay watershed. *Remote Sens. Environ.*
- Thomlinson, J.R., Bolstad, P.V., Cohen, W.B., 1999. Coordinating methodologies for scaling landcover classifications from site-specific to global. *Remote Sens Environ.* 70, 16–28.
- Thornthwaite, C.W., 1948. An Approach toward a Rational Classification of Climate. *Geogr. Rev.* 38, 1 (55-94).
- Tittonell, P., El Mujtar, V., Felix, G., Kebede, Y., Laborda, L., Luján Soto, R., de Vente, J., 2022. Regenerative agriculture—agroecology without politics? *Frontiers in Sustainable Food Systems* 6.
- Tóth, G., Jones, A., Montanarella, L. 2013. The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. *Environ. Monit. Assess.*, 185(9), 7409-7425.
- Tóth, G., Song, X., Hermann, T., Tóth, B. 2016. Hierarchical and multi-scale pedoclimatic zonation. *iSQAPER Project Deliverable 2.1*, 125 pp.
- UC Business Analytics R Programming Guide, 2018. Gradient Boosting Machines. http://uc-r.github.io/gbm_regression.
- Ujaval, C., 2020. Extracting Time Series using Google Earth Engine. <https://spatialthoughts.com/2020/04/13/extracting-time-series-ee/>
- University of East Anglia Climatic Research Unit; Harris, I.C.; Jones, P.D.; Osborn, T. 2020. CRU TS4.04: Climatic Research Unit (CRU) Time-Series (TS) version 4.04 of high-resolution gridded data of month-by-month

- variation in climate (Jan. 1901- Dec. 2019). Centre for Environmental Data Analysis.
<https://catalogue.ceda.ac.uk/uuid/89e1e34ec3554dc98594a5732622bce9>
- van Deventer, A.P., Ward A.D., Gowda P.H., Lyon J.G., 1997. Using Thematic Mapper data to identify contrasting soil plains and tillage practices. *Photogramm. Eng. Remote Sens.* 63 (1):87-9
- Van Orshoven, J., Maes, J., Vereecken, H., Feyen, J., Didal, R., 1988. A structured database of Belgian soil profile data. *Pédologie.* 38, 191–206.
- van Wesemael, B., Chartin, C., Wiesmeier, M., von Lützw, M., Hobley, E., Carnol, M., Krüger, I., Campion, M., Roisin, C., Hennart, S., Kögel-Knabner, I. 2019. An indicator for organic matter dynamics in temperate agricultural soils. *Agric. Ecosyst. Environ.*, 274, 62-75.
<https://doi.org/10.1016/j.agee.2019.01.005>
- Van Wesemael, B., Paustian, K., Andrén, O., Cerri, C.E.P., Dodd, M., Etchevers, J., Goidts, E., Grace, P., Kätterer, T., McConkey, B.G., Ogle, S., Pan, G., Siebner, C. 2011. How can soil monitoring networks be used to improve predictions of organic carbon pool dynamics and CO₂ fluxes in agricultural soils? *Plant and Soil*, 338, 247–259.
- van Wesemael, B., Paustian, K., Meersmans, J., Easter, M., Schlesinger, W.H. (2010). Agricultural management explains historic changes in regional soil carbon stocks. *PNAS.* 107(33), 14926-14930.
<https://doi.org/10.1073/pnas.1002592107>
- Vandenbergh C., Marcoen J.M., Sohier C., Degré A., Hendrickx C., Paulus P., 2009. Monitoring networks and modelling systems for assessing effectiveness of the EU Nitrates Directive Action Programmes: Approach by the Walloon Region (Belgium). Dossier GRENeRA 09-08 26p.. In Marcoen J.M., Lambert R., Vandenbergh C., De Toffoli M., Benoit J., Deneufbourg M., 2010. Programme de gestion durable de l'azote en agriculture wallonne – Rapport d'activités annuel intermédiaire 2009 des membres scientifiques de la Structure d'encadrement Nitrawal. Université de Liège Gembloux AgroBio Tech et Université catholique de Louvain, 62p + annexes
- Varvel, G.E., Wilhelm, W.W. (2010). Long-Term Soil Organic Carbon as Affected by Tillage and Cropping Systems. *Soil Sci. Soc. Am. J.*, 74(4), 915-921. <https://doi.org/10.2136/sssaj2009.0362>
- Vaudour, E., Gomez, C., Fouad, Y., Lagacherie, P., 2019. Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. *Remote Sens. Environ.* 223, 21–33.
<https://doi.org/10.1016/j.rse.2019.01.006>.
- Villarino, S., Studdert, G., Baldassini, P., Cendoya, M., Ciuffoli, L., Mastrángelo, M. and Piñeiro, G., 2017. Deforestation impacts on soil organic carbon stocks in the Semiarid Chaco Region, Argentina. *Sci. Total Environ.* 575, pp.1056-1065.
- Viscarra Rossel, R.A., Webster, R., Bui, E.N., Baldock, J.A., 2014. Baseline map of organic carbon in Australian soil to support national carbon

accounting and monitoring under climate change. *Glob. Chang. Biol.* 20:2953-2970.

Wang, C., Pan, X., Liu, Y., Li, Y., Shi, R., Zhou, R., Xie, X., 2013. Alleviating moisture effects on remote sensing estimation of Crop Residue Cover. *Agronomy Journal* 105, 967–976.

Wang, X., Blesh, J., Rao, P., Paliwal, A., Umashaanker, M., Jain, M., 2023. Mapping cover crop species in southeastern Michigan using Sentinel-2 satellite data and Google Earth Engine. *Front. Artif. Intell.* 6: 1035502.

Wang, X., Qi, J.-Y., Zhang, X.-Z., Li, S.-S., Virk, A. L., Zhao, X., Xiao, X.-P., Zhang, H.-L. 2019. Effects of tillage and residue management on soil aggregates and associated carbon storage in a double paddy cropping system. *Soil Tillage Res.*, 194, 104339. DOI: 10.1016/j.still.2019.104339.

Watson, S. J., Luck, G. W., Spooner, P. G., Watson, D. M. 2013. Land-use change: incorporating the frequency, sequence, time span, and magnitude of changes into ecological research. *Ecol. Appl.*, 23(8), 1417-1428.

Wei, X., Shao, M., Gale, W., Li, L. 2014. Global pattern of soil carbon losses due to the conversion of forests to agricultural land. *Scientific Reports*, 4, 4062.

Weiss, M., Jacob, F., Duveiller, G. 2020. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* 236, 111402.

West, T.O., Post, W.M., 2002. Soil Organic Carbon Sequestration Rates by Tillage and Crop Rotation. *Soil Sci. Soc. Am. J.* 66(6), 1930–1946.

Wiesmeier, M., von Lützw, M., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., Schilling, B., Kögel-Knabner, I. 2015. Land use effects on organic carbon storage in soils of Bavaria: The importance of soil types. *Soil Tillage Res.*, 146, 296–302. <https://doi.org/10.1016/j.still.2014.10.003>

Wiesmeier, M., Lungu, M., Cerbari, V., Boincean, B., Hübner, R., Kögel-Knabner, I., 2018. Rebuilding soil carbon in degraded steppe soils of Eastern Europe: The importance of windbreaks and improved cropland management. *Land Degrad Dev.* 29, 875– 883. <https://doi.org/10.1002/ldr.2902>.

Wiesmeier, M., Poeplau, C., Sierra, C.A., Maier, H., Frühauf, C., Hübner, R., Kühnel, A., Spörlein, P., Geuß, U., Hangen, E., Schilling, B., von Lützw, M., Kögel-Knabner, I. 2016. Projected loss of soil organic carbon in temperate agricultural soils in the 21st century: effects of climate change and carbon input trends. *Scientific Reports*, 6, Article 32525. <https://doi.org/10.1038/srep32525>

Wiesmeier, M., Urbanski, L., Hobbey, E., Lang, B., von Lützw, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H.J., Kögel-Knabner, I., 2019. Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales. *Geoderma*, 333, 149–162. <https://doi.org/10.1016/j.geoderma.2018.07.026>.

- Wiltshire, S., Beckage, B. 2023. Integrating climate change into projections of soil carbon sequestration from regenerative agriculture. *PLoS Clim.*, 2(3), e0000130.
- Xie, E., Zhang, X., Lu, F., Peng, Y., Chen, J., Zhao, Y. 2022. Integration of a process-based model into the digital soil mapping improves the space-time soil organic carbon modelling in intensively human-impacted area. *Geoderma*, 409, 115599. DOI: 10.1016/j.geoderma.2021.115599.
- Xu, L., He, N. P., Yu, G. R., Wen, D., Gao, Y., He, H. L. 2015. Differences in pedotransfer functions of bulk density lead to high uncertainty in soil organic carbon estimation at regional scales: Evidence from Chinese terrestrial ecosystems. *J. Geophys. Res. Biogeosci.*, 120(11), 2215-2226.
- Yan, L., Roy, D.P., 2020. Spatially and temporally complete landsat reflectance time series modelling: The fill-and-fit approach. *Remote Sens Environ.* 241, 111718.
- Yang, X., Wang, M., Huang, Y., & Wang, Y. (2002). A one-compartment model to study soil carbon decomposition rate at equilibrium situation. *Ecological Modelling*, 151(1), 63-73.
- Yigini, Y., Panagos, P. 2016. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. *Sci. Total Environ.*, 557-558, 838-850.
- Yokozawa, M., Shirato, Y., Sakamoto, T., Yonemura, S., Nakai, M., Ohkura, T. 2010. Use of the RothC model to estimate the carbon sequestration potential of organic matter application in Japanese arable soils. *Soil Sci. Plant Nutr.* 56, 168-176.
- Zhang, L., Heuvelink, G.B.M., Mulder, V.L., Chen, S., Deng, X., Yang, L. 2024. Using process-oriented model output to enhance machine learning-based soil organic carbon prediction in space and time. *Sci. Total Environ.* 922, 170778.
- Zheng, B., Campbell, J.B., Serbin, G., Galbraith, J.M., 2014. Remote sensing of crop residue and tillage practices: Present capabilities and future prospects. *Soil Tillage Res.* 138, 26–34.
- Zhou, T., Geng, Y., Chen, J., Pan, J., Haase, D., Lausch, A., 2020. High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms. *Sci. Total Environ.* 729, 138244. <https://doi.org/10.1016/j.scitotenv.2020.138244>.
- Zhou, Y., Chartin, C., Van Oost, K., van Wesemael, B., 2022. High-resolution soil organic carbon mapping at the field scale in southern Belgium (Wallonia). *Geoderma*. 422, 115929. <https://doi.org/10.1016/j.geoderma.2022.115929>
- Zhou, Y., Chen, S.C., Zhu, A.X., Hu, B.F., Shi, Z., Li, Y., 2021. Revealing the scale-and location-specific controlling factors of soil organic carbon in Tibet. *Geoderma*, 382, 114713. <https://doi.org/10.1016/j.geoderma.2020.114713>

Zhou, Y., Ferdinand, M., van Wesemael, J., Dvorakova, K., Baret, P., Van Oost, K., van Wesemael, B. A framework for mapping conservation agricultural fields using time-series optical and radar imagery. (Under review).

Zomer, R.J., Bossio, D.A., Sommer, R. Verchot, L. Global Sequestration Potential of Increased Organic Carbon in Cropland Soils. *Sci Rep* 7, 15554 (2017).

Zribi, M., & Dechambre, M. 2003. A new empirical model to retrieve soil moisture and roughness from C-band radar data. *Remote Sensing of Environment*, 84(1), 42-52.

List of Figures

Fig. 1.1. Levers associated with agricultural practices that may influence SOC stocks: (1) increasing primary production (e.g. crop rotations, agroforestry, cover crops), (2) increasing biomass return to soil (crop residue return), (3) importing organic wastes to soil (manures, composts..), (4) avoiding fires, (5) grassland management (fertilization, grazing), (6) decreasing biodegradation and mineralisation rates (no tillage, water management), (7) decreasing erosion rates. (Chenu et al., 2019).....	2
Fig. 1.2. Schematic diagram of DSM in cropland (Huang et al., 2022)	4
Fig. 1.3. Percentage change in soil organic carbon (SOC) due to land management practices in cropland (Beillouin et al., 2023).....	6
Fig. 1.4. (a) Percentage of models in different classes; (b)typical spatial and temporal scales of the model classes. The lines show the spatial and temporal ranges where 50% of the models in each class fall. (Manzoni & Porporato, 2009).....	7
Fig. 1.5. An example of a single-pool soil organic matter dynamic model and the connection to hypothetical data used to formulate (A), calibrate (B), drive (C), and evaluate (D) model functions. (Campbell & Paustian, 2015).....	8
Fig. 1.6. Components of a soil measurement/monitoring (M), reporting (R) and verification(V) framework (Smith et al., 2020).....	10
Fig. 1.7. Predict SOC spatial distribution at t=0	14
Fig. 1.8. Simulate SOC change over time using RothC model with long-term experiment (LTE) data.	15
Fig. 1.9. Prepare large-scale climate and management data using remote sensing, e.g. time-series NDVI could be used to analyze presence of cover crop.....	15
Fig. 1.10. Predict SOC spatial distribution at t=1.	16
Fig. 1.11. Calculate SOC Sequestration between different scenarios.....	16
Fig. 1.12. Study area for each chapter of the thesis.	17
Fig. 2.1. Digital elevation model of Wallonia and locations of 497 soil samples (A) (70% calibration samples in white and 30% validation samples in red) and 21,551 external validation data (B).....	24
Fig. 2.2. Workflow of extracting time-series NDVI values in the Google Earth Engine.....	27
Fig. 2.3. Typical example of the NDVI time series in a single field used to illustrate the identification of the frequency of cover crops (n=2). 28	28
Fig. 2.4. Maps of Covariates (A)OC input (B)NDVI (C)Precipitation (D)Clay content. Red line: a transect selected for section 2.2.4	33
Fig. 2.5. Predicted vs. observed SOC contents of validation dataset (n = 149) (red line is the 1:1 line). The error bars are the standard deviations obtained by the bootstrapping.	37

- Fig. 2.6. Predicted vs. observed SOC contents of external validation dataset (n = 21,551) (red line is the 1:1 line) and the histogram of the residues.....37
- Fig. 2.7. The raincloud plot of the covariates' importance in 50 simulations, including an illustration of importance distribution (the cloud) with the raw data (the rain).38
- Fig. 2.8. Original SOC signal, residue and four IMFs of EMD, and correlation coefficients with covariates. (IMF: Intrinsic mode functions; EMD: Empirical mode decomposition)39
- Fig. 2.9. Spatial distribution of (A) mean SOC content and relative 90% prediction limit range (B), zoomed to a smaller region (C) with its interquartile range (D), and land use map (E). The white areas are forests or built-up areas.41
- Fig. 2.10. Comparison of zoomed SOC content maps from our work (A1) (B1) and the Walloon SOC map (A2) (B2).....44
- Fig. 2.11. The distribution of SOC content ($\text{g}\cdot\text{kg}^{-1}$) (A) in an area with recent land use change (B): land use in 1990, (C): land use in 2000.47
- Fig. 2.12. Map of different aggregate stability classes (i.e. highly unstable, unstable and stable) and histogram of misclassification potential (The probability of not belonging to the current class among the 50 predicted results, negative values indicate downgrading).....48
- Fig. 3.1. Digital elevation model and agro-geographical regions of the Walloon region, the bar plots display the number of conservation (green bar) and conventional (blue bar) agriculture fields in the survey.60
- Fig. 3.2. Examples of identifying cover crops and bare soil based on long-term series NDVI and NBR2 values during 2015 to 2020. Field A: This field was fallow between the two annual crops from 2016 to 2017 (Inter2) and consisted of a cover crop and bare soil between the two annual crops from 2017 to 2018 (Inter3), while the next year (inter4) was bare soil. Field B: Cover crop was planted between the annual crops of 2015 and 2016 (inter1).69
- Fig. 3.3 Workflow for building tillage models and predicting non-inversion/inversion tillage practice under three scenarios (Here: in orange tables of Step 3, Tillage represents the actual tillage practices and Tillage' represents predicted tillage practices using Model I & Model II. The procedure of random forest and prediction in Step3 used 10-fold cross validation; here we simplify this part in this figure).73
- Fig. 3.4. Grouped boxplots with significance notation under Wilcoxon test, comparing the levels of two indicators related with species diversification under two farming system in Walloon region during 2015-2020. (ns: not statistically significant ($P > 0.05$), *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$);. (Here: CA: Conservation

	Agriculture; TA: Conventional Agriculture). The blue numbers in each subplot indicate P value of Wilcoxon test.....	77
Fig. 3.5.	Grouped boxplots with significance notation under Wilcoxon test, comparing the levels of 4 indicators related with soil cover under two farming system in Walloon region during 2015-2020 (ns: not statistically significant ($P > 0.05$), *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$);. (Here: CA: Conservation Agriculture; TA: Conventional Agriculture). The blue numbers in each subplot indicate P value of Wilcoxon test.....	78
Fig. 3.6.	Importance of covariates in predicting tillage practices under Model I: Without cover crops and Model II: With cover crops.	79
Fig. 3.7.	(A) Model accuracy evolution for different number of covariates under recursive feature elimination and (B) importance of covariates. (Tillage: the tillage practice; N_crop: number of crop types; P_BS: bare soil period; P_fallow: fallow period) ...	80
Fig. 3.8.	Classified map (A) and enlarged views in representative regions are presented in (B1) and (B2), with the two most important indicators, tillage practice and the number of crop types (for the period 2015-2020), only croplands are mapped leaving grassland, forest and built up areas in white.	82
Fig. 4.2.	Effect of replacing measured boundary conditions with remote sensing derived one, which calculated based on the difference under two models.	108
Fig. 4.3.	Comparison of recorded cover crop calendar and	111
Fig. 5.1.	(A) Location and extent of the study area, where orange grids represent 0.125° ORCHIDEE grids. (B) Distribution of fields within the study area and their initial SOC stock values (from section 5.2.3.1 Soil data).	120
Fig. 5.2.	(A) Distribution of conservation and conventional agriculture fields; for zoomed region (B) Field category, tillage practice, number of different crops and annual presence of cover crops are shown (Source: Chapter 3).	125
Fig. 5.3.	(A) European Pedoclimate zones (B) Belgian Soil Association groups	129
Fig. 5.4.	SOC stock changes ($t\ ha^{-1}$) during 2015-2024 for model runs at different scales(A) Per field (B) Belgian Soil association groups and (C) European Pedoclimate zones (see Fig. 5.1 for details) ...	132
Fig. 5.5.	Density plot of SOC stock changes in conventional and conservation agriculture systems from 2015 to 2024.....	135
Fig. 5.6.	Density plot of SOC stock changes per field under different management scenarios from 2015 to 2024, dash line represents the uncertainty value obtained from section 4.3.2 (BAU: Business-as-usual; RT: reduced tillage; CC: Long duration cover crop).	137
Fig. 5.7.	Compare SOC stock changes ($t\ ha^{-1}$) between Business-as-Usual and Long-term Cover crop Scenarios during 2015-2024 for (A) two	

example fields (B) Three example fields (C) Averaged value for entire region and (D) Two soil associations. (Region 31: Loamy Soils with textural B horizon (well-drained); Region 35: Loamy Soils with mottled textural B horizon) 141

List of Tables

Table 2.1. Covariates used for modeling SOC contents. (R: Resolution).....	29
Table 2.2. Summary descriptive statistics of measured SOC content ($\text{g}\cdot\text{kg}^{-1}$) in calibration and validation datasets.....	36
Table 2.3. Crop types and total OC input (include CC) in the field over a five-year period. (CC indicates the existence of a cover crop after the main crop).....	45
Table 3.1. Band wavelengths of Landsat-7, Landsat-8, and Sentinel-2 satellite sensors (W: Wavelength; R: Resolution).	65
Table 3.2. Composition of the confusion matrix (NN represents the number of fields considered conventional by both measured and predicted data, CC represents the number of fields considered conservation by both measured and predicted data, CN represents the number of fields that are measured as conservation but predicted as conventional, NC represents the number of fields that are measured as conventional but predicted as conservation.)	76
Table 3.3. Confusion matrix of cover crop classification (Where CC means Cover Crop)	77
Table 3.4. Confusion matrix of non-inversion/inversion tillage mode	79
Table 3.5. Confusion matrix of field classification model.....	81
Table 4.1. Summary of the simulations under different input data combinations (Green cell represents in-situ data;.....	103
Table 4.2. Predicted final SOC stock and	108
Table 5.1. Boundary conditions for RothC model and their source	122
Table 5.2. Utilized agricultural area (UUA), agricultural land proportion and annual input of farmyard manure (FYM, t C ha^{-1}) by province..	127
Table 5.3. Ratio of Planting Cover Crop in Study Area.	130
Table 5.4. SOC stock changes during 2015-2024 for 50,655.52 ha of agricultural fields in the Walloon region of Belgium.	133
Table 5.5. Soil organic carbon (SOC) stock changes during 2015-2024 for 44,004.19 ha of conventional agricultural fields in the Walloon region of Belgium, calculated based on per field.	138