Carbon cycle as affected by soil erosion in Europe

Arthur Fendrich

université PARIS-SACLAY

Carbon cycle as affected by soil erosion in Europe

Le cycle du carbone affecté par l'érosion des sols en Europe

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 581: Agriculture, alimentation, biologie, environnement, santé (ABIES)
Spécialité de doctorat: Sciences de l'environnement
Graduate School: Biosphera. Référent: AgroParisTech

Thèse préparée dans l'UMR **SAD-APT** (Université Paris-Saclay, INRAE, AgroParisTech), sous la direction de **Philippe MARTIN**, Professeur, la co-direction de **Panos PANAGOS**, Chercheur (HDR), le co-encadrement de **Philippe CIAIS**, Chercheur, et de **Emanuele LUGATO**, Chercheur

Thèse soutenue à Paris-Saclay, le 05 octobre 2023, par

Arthur Nicolaus FENDRICH

Composition du jury

Membres du jury avec voix délibérative

Christine ALEWELL Professeure, University of Basel (Suisse)	Présidente
Julia PONGRATZ Professeure, LMU (Allemagne)	Rapporteur et Examinatrice
Pierre REGNIER Professeur, Université Libre de Bruxelles (Belgique)	Rapporteur et Examinateur
Olivier CERDAN Chercheur, BRGM	Examinateur
Timothy QUINE Professeur, University of Exeter (Royaume-Uni)	Examinateur

THÈSE DE DOCTORAT
NNT : 2023UPASB048

université PARIS-SACLAY
ÉCOLE DOCTORALE
Agriculture, alimentation, biologie, environnement, santé (ABIES)

Titre: Le cycle du carbone affecté par l'érosion des sols en Europe

Mots clés: Erosion des sols, cycle du carbone, modélisation du climat

Résumé: Si l'importance de l'impact de l'érosion des sols sur le cycle du carbone est désormais bien reconnue dans la littérature, la représentation de ce processus dans les modèles de surface terrestre doit encore être améliorée, en particulier à grande échelle. De même, l'adoption de cultures de couverture en tant qu'alternative de gestion durable des sols a suscité un intérêt politique croissant ces dernières années, malgré des impacts encore non quantifiés sur le carbone organique du sol et l'érosion. Dans cette thèse, nous avons abordé ces problématiques de trois manières.
Premièrement, nous avons amélioré le modèle CE-DYNAM, qui couple le mouvement latéral des particules du sol à un émulateur du cycle du carbone organique du sol, permettant son application à des domaines plus larges et à des résolutions spatiales plus fines. Ce travail a également introduit une procédure de calibration pour CE-DYNAM afin de rendre ses résultats plus réalistes et plus cohérents avec les observations de décharge de sédiments dans les stations fluviales.
Deuxièmement, nous avons développé un modèle de désagrégation statistique afin générer la première carte spatialement explicite des cultures de couverture pour l'Europe. Les résultats ont montré que les données des radars à synthèse d'ouverture pouvaient contribuer à la détection à distance des cultures de couverture et ont mis en évidence l'importance de mettre à la disposition des chercheurs des données sur la gestion des terres. Cette carte devrait être utilisée par les chercheurs et les praticiens dans le cadre de travaux ultérieurs.
Enfin, nous avons combiné les développements des chapitres précédents pour modéliser l'impact des scénarios politiques concernant les cultures de couverture sur le cycle du carbone. Nos résultats indiquent que les cultures de couverture peuvent simultanément augmenter le stockage de carbone organique dans le sol tout en réduisant l'exportation de carbone organique particulaire vers les océans, mais qu'il existe une limite à ces deux effets. Le pic d'accumulation de carbone organique dans le sol se produit avant la stabilisation de l'exportation vers les océans, indiquant un délai qui peut s'expliquer par le temps nécessaire aux particules pour se déplacer dans le paysage.

Title: Carbon cycle as affected by soil erosion in Europe

Keywords: Soil erosion, carbon cycle, climate modeling

Abstract: While the importance of the impact of soil erosion on the carbon cycle is now well-recognized in the literature, the representation of this process in land surface models still needs to be better developed, especially at a large scale. Similarly, adopting cover crops as a sustainable soil management alternative has had growing policy interest in recent years despite the still unquantified impacts on soil organic carbon and erosion. In this thesis, we addressed these problems in three ways.
First, we advanced the development of the CE-DYNAM model, which couples the lateral movement of soil particles to an emulator of the soil organic carbon cycle, enabling its application at larger domains and finer spatial resolutions. This work also introduced a calibration procedure for CE-DYNAM to make its results more realistic and coherent with sediment discharge observations in river stations.
Second, we developed a statistical disaggregation model to generate the first spatially explicit map of cover crops for Europe. This work indicated that synthetic aperture radar data could assist the remote detection of cover crops, and highlighted the importance of making land management data publicly available for researchers. The map is expected to be used by researchers and practitioners in further works.
Finally, we combined the developments of the previous chapters to model the impact of policy scenarios for cover crops on the carbon cycle. Our results indicate that cover crops can simultaneously increase soil organic carbon storage while reducing particulate organic carbon export to the oceans, but a limit exists for both effects. The peak in soil organic carbon accumulation happens before the stabilization of ocean export, indicating a delay that can be explained by the time taken by particles to travel across the landscape.

This PhD was funded and supported by the European Commission (grant no. 35403) and by the CLAND project.

Acknowledgements

During my PhD, I once pictured a fictional dialogue of two nobles, centuries apart, one saying that « I am among those who think that science has great beauty » and the other replying that « beauty walks a razor's edge, someday I'll make it mine ». Writing this thesis felt no different than traveling such a rough path. The awe of dedicating a part of my lifetime to comprehend a natural phenomenon was sometimes confusing when seemingly simple things disguised a complexity that could only be seen through the lens of a scientific magnifier. This period had many unforgettable moments that helped me to stay on track, and here I would like to express my gratitude to many of the meaningful relationships built and strengthened during these moments by dedicating this work:

To my advisors Philippe Martin, Panos Panagos, Philippe Ciais, Emanuele Lugato, Marco Carozzi, and Bertrand Guenet for caring about the work and its development, the many insightful discussions, and all the support and encouragement. Having such a high-level and experienced group of advisors gave me extra safety for writing the thesis, and I am really grateful for that. This work would not have been possible without each of your individual contributions, and I hope to carry your attitude and enthusiasm with me during my post-PhD life and career. I also thank Victoria Naipal for sharing her codes and ideas about the CE-DYNAM model.

To Julia Pongratz, Pierre Reigner, Christine Alewell, Timothy Quine and Olivier Cerdan for agreeing to evaluate my thesis and thinking of comments that will improve the quality of the work done. Likewise, to Pasquale Borrelli and Kristof Van Oost for monitoring the progress of the work over the years and for methodological and theoretical remarks that allowed me to stay on its main subject.

To all my colleagues of the BIOGEO group (LSCE/France), the soil team (JRC/Italy), and the friends I made in both countries. I could learn many subjects in our meetings and had much fun in the informal exchanges during coffee time. A special thank you to my CDP colleagues Francis Matthews, Julia Köninger, and Anna Muntwyler, and to those that dissolved the boundaries between professional and personal life: to Maëva Labouyrie for the relaxing pizza moments and the controlled stress outbursts; to Elyse Van Eynde and Ismael de Bauw for the company during office hours and for teaching me so much about work and life; to Timo Breure and Anne Maréchal for the long but funny bus trips; to Yantong Lin for all your knowledge that comforted me many times during the thesis, I hope I haven't hurt your ears.

To Caroline, for all the love and emotional support. This thesis demanded very complex arrangements, and it never ceased to amaze me how you overcame them with so much energy, joy, lightness, and confidence. Thank you for bringing so much color to my life.

To my family - my mother Silvia, my father Alexandre, and my sister Lorena - for the love and guidance in my life decisions. Thank you very much!

And what did you hear, my blue-eyed son?
And what did you hear, my darling young one?
I heard the sound of a thunder that roared out a warning
I heard the roar of a wave that could drown the whole world
I heard one hundred drummers whose hands were a-blazing
I heard ten-thousand whispering and nobody listening
I heard one person starve, I heard many people laughing
I heard the song of a poet who died in the gutter
I heard the sound of a clown who cried in the alley
(...) It's a hard rain's a-going to fall

— Bob Dylan

Contents

List of abbreviations

Throughout this thesis, the reader will encounter the following

Abbreviations:

- AUC - Area Under the Curve
- C - Carbon
- CAP - Common Agricultural Policy
- CCs - Cover crops
- CR - Cross-polarization Ratio
- CSR - Compressed Sparse Row
- EFAs - Ecological Focus Areas
- ESDAC - European Soil Data Centre
- ETD - Erosion, Transport and Deposition
- EU - European Union
- FSS - Farm Structure Survey
- GSAA - Geospatial Aid Application
- IACS - Integrated Administration and Control System
- LPIS - Land Parcel Identification System
- LSM - Land surface model
- LUCAS - Land Use/Cover Area frame statistical Survey
- ME - Model efficiency
- MSs - Member States
- NUTS - Nomenclature of Territorial Units for Statistics
- PFT - Plant functional types
- POC - Particulate organic carbon
- ROC - Receiver Operating Characteristic
- RPG - Registre Parcellaire Graphique
- (R)USLE - (Revised) Universal Soil Loss Equation
- SAR - Synthetic Aperture Radar
- SOC - Soil organic carbon
- UK - United Kingdom

Units:

- Gg - Giga grams (= 10^9 grams)
- Tg - Tera grams (= 10^{12} grams)
- Pg - Peta grams (= 10^{15} grams)
- t/ha/year - Tonnes per hectare per year

1 - General introduction

1.1 . Soil erosion and climate change

Human activities have been significantly affecting our climate system for the past 12,000 years (Brooke, 2018), causing impacts that pose risks to natural systems. Increases in sea level rise rates, the number of heavy precipitation events, ocean temperatures, and the frequency and duration of heat waves are examples of impacts that are at least likely to have happened due to human alterations (IPCC, 2013a, 2014). Because of these and other consequences, climate change is among the major environmental challenges for humanity in the 21st century (Jones et al., 2012; King, 2004; Lorey, 2003; United Nations Environment Programme, 2012)

Climate change is formally defined as the alterations in global atmospheric composition that can be directly or indirectly attributed to human activity, therefore additional to the natural variability observed (United Nations, 1992). The emissions from human activities have intensified since the First Industrial Revolution (nearly 1760), leading to an ever-increasing concentration of greenhouse gases in the atmosphere. Past data show that fossil fuel emissions increased from zero in pre-industrial years to 9.5 PgC year^{-1} in the 2010s (IPCC, 2013b). Several recent extreme events show that the effects of climate change are already being perceived throughout the globe, which highlights the need for actions and for a deep understanding of its physical processes (Walsh et al., 2018; Fosu et al., 2018; Herring et al., 2015; Zhou et al., 2018; CarbonBrief, 2021).

In the context of the Earth's climate system, soils are an essential natural resource. Many of the ecosystem services provided by soils are fundamental to supporting human livelihood and are affected by climate change. These include the production of food, freshwater, energy, and habitat provision for biodiversity. The human demand for these services, and consequently the pressure on soils, tends to increase with the projected global population growth (IPCC, 2019). Furthermore, world soils contain 1500-2400 PgC, more than the atmosphere (589 PgC) and the surface ocean (900 PgC) together (Ciais et al., 2013). Therefore, even small disturbances in soil carbon pools can substantially impact their emissions to the atmosphere. For example, more sustainable management actions, such as retaining residues in cropping systems, can increase the soil's organic carbon content and help sequester carbon

from the atmosphere (Robert, 2001), and mitigate a major soil threat, soil erosion.

Accelerated soil erosion is the most widespread process of soil degradation caused by agricultural activities (Robert, 2005), a list that includes compaction, nutrient depletion, breakdown of aggregates, and others. According to Hillel (2004), the physical process of soil erosion can be described in three steps: i) First, particles are detached from the soil. The most vulnerable particles are those of the soil surface, which is the most fertile part of the profile for the large amounts of organic matter (humus) and microorganisms. The loss of fertility caused by the detachment of topsoil particles usually demands the use of chemical fertilizers, which leads to an increasing threat to groundwater pollution. ii) Second, the detached particles are transported through the landscape; and iii) Third, particles are deposited in a different place where they originally belonged. Alternatively, authors sometimes include the breakdown of macroaggregates as an intermediate process between the first and the second above (Lal, 2005). Soil erosion can be triggered by water or wind, the former being the most common. The schema of Figure 1.1 represents the three steps described by Hillel (2004).

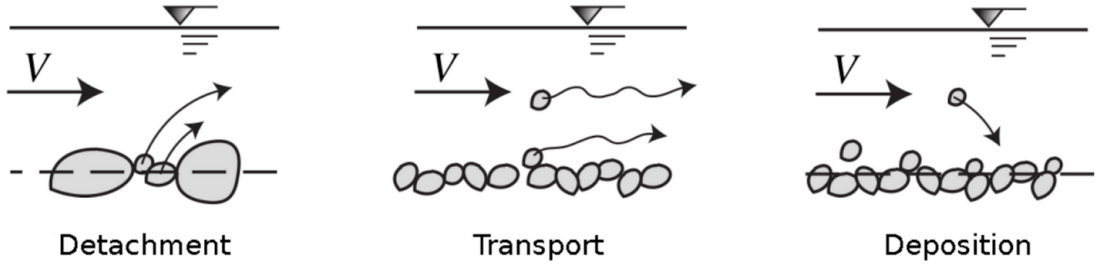


Figure 1.1: The three steps of soil erosion by water. V represents water flow velocity. Adapted from: Julien (2010).

The flow of particles through the landscape is guided by gravity and follows the geomorphological characteristics of the terrain. Consequently, soil erosion can have on-site impacts, such as the loss of productivity, and off-site impacts, such as the eutrophication of rivers and reservoirs. For this reason, soil erosion is understood as a key transport mechanism to carry the environmental effects of agricultural chemicals to farther areas (Hillel, 2004). Yearly, 20 - 30 Gt of soils worldwide are lost because of water erosion, with the rates per unit of area varying according to the climatic zone. In temperate regions, average values are usually up to 10 t ha^{-1} $year^{-1}$ but may reach 20 t ha^{-1} $year^{-1}$ in hilly cropland areas. In tropical regions,

which concentrate the world's highest erosion rates, values often range from 10 to 20 t ha^{-1} $year^{-1}$, and may be as high as 50 t ha^{-1} $year^{-1}$ in more extreme cases (FAO, 2014; van Oost et al., 2007). These values are supported by the more recent results of Borrelli et al. (2018a), which reported an average soil loss of 35 Gt in 2001 and 35.9 Gt in 2012, with the highest area-specific rates in South America, Africa, and Asia, respectively.

The convergence of erosion predictions in the literature implies that regardless of the database used, the values are much higher than the soil formation rates (Julien, 2010). These are estimated to be generally lower than 1 t ha^{-1} $year^{-1}$, with a median of 0.15 t ha^{-1} $year^{-1}$ (FAO, 2014). Such an imbalance between soil formation and loss means world soils could be treated as a non-renewable instead of a renewable natural resource. Besides, in the long-term scale of decades or centuries, these erosion rates are likely to directly reduce crop yield due to lower water holding capacity and root space accommodation (FAO, 2014).

The link between soil erosion and the carbon cycle is shown in Figure 1.2. On-site, all mechanisms are understood to lead to increased atmospheric emissions. These include the removal of clay and soil organic carbon, the increase in mineralization rates by alterations in soil moisture, the subsurface exposure, and the breakdown of aggregates. Off-site, some of the effects, such as the transport to other landscape elements, such as floodplains and aquatic systems, could lead to the protection and sequestration of carbon. The intensity of each of these effects has led to a long ongoing scientific debate to understand if soil erosion corresponds to a source or a sink of atmospheric carbon, with results pointing to different directions.

On the one hand, Stallard (1998) constructed hundreds of simulations to span multiple possible scenarios and concluded that the thesis of erosion being a C sink is plausible. More recently, van Oost et al. (2007) revisited this hypothesis in an assessment that found that soil erosion can be a global sink of C with a magnitude of 1 PgC $year^{-1}$. On the other hand, Lal (2003) argues that soil erosion is a C source, with a magnitude comparable to that of van Oost et al. (2007). Such an argument was also reinforced by the continental assessment of Lugato et al. (2018). In the future, other results can come from further developments in representing the erosion process in land carbon models.

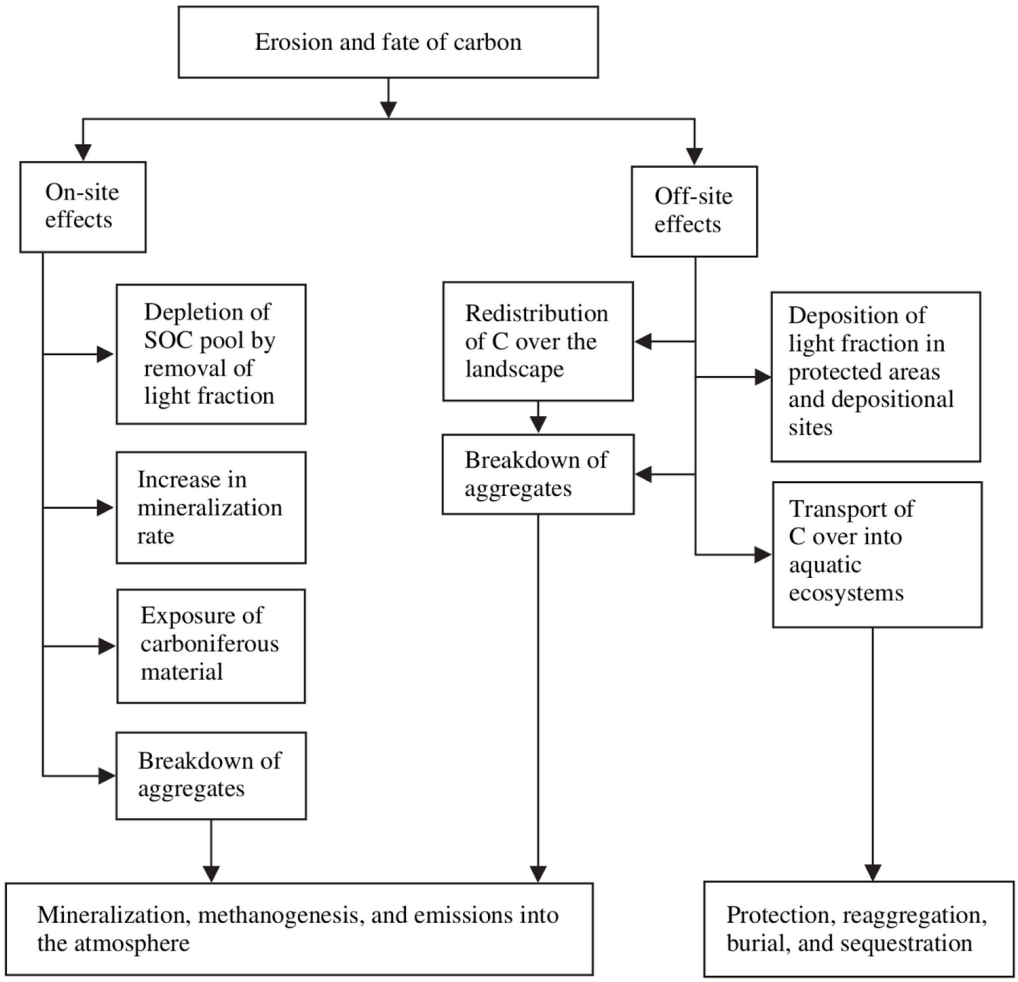


Figure 1.2: The physical mechanisms that link the carbon cycle to erosion, transport and deposition processes. Source: Lal (2005).

1.2 . Modeling erosion, transport, and deposition

Although mentions of soil erosion as an agricultural issue date back to the 1810s, it was only in the 1920s that soil erosion began to be recognized as an important problem for soils in the United States. The effort of scientists to gather public support to fight the "menace to the national welfare" precedes the so-called "golden years" of conservation research (1930-1942), when agricultural plots for erosion research were installed across the country. In the meantime, the great dust storm of 1934 had already increased public awareness to conserve soil resources and from the 1940s on, different numerical approaches tried to take advantage

of the data gathered so far to explain and predict erosion based on various factors, such as land slope, slope length, storm rainfall energy, and others. The progress was harmed by the advent of World War II, which led to an abrupt interruption of erosion research. Even so, the Musgrave equation proposed in 1947 combined many models developed so far, such as those developed by Zingg, Smith and Browning (Laflen and Flanagan, 2013), and can be seen as one of the first attempts to obtain a general model for soil erosion (Laflen, 2003). Its formulation was:

$$A' = \left(\frac{P_{30}}{1.25}\right)^{1.75} K' \left(\frac{L}{72}\right)^{0.35} \left(\frac{S}{10}\right)^{1.35} C^*$$

with A' being the soil loss (inches of depth), P_{30} being the maximum precipitation amount in 30 minutes of storm (inches), K' being the Musgrave soil erodibility factor (inches/year), L being the slope length (feet), S being the slope gradieng (%), and C^* being the vegetal cover factor (Laflen and Flanagan, 2013).

The Musgrave equation was widely adopted during the following years. However, later works showed that some of its assumptions could not be justified and that its rainfall factor needed to be revised for different situations in the United States. Therefore, it was the moment to reevaluate the Musgrave equation, and in the 1950s, workshops were held to extract knowledge from the data collected so far. After compiling data for 10,000 plots across the country, it was in 1959 that Wischmeier (1959) published his work "A rainfall erosion index for a Universal Soil Loss Equation", which represents the first time that the term Universal Soil Loss Equation (USLE) was formally mentioned. However, despite the suggestive name, W. H. Wischmeier later declared, in 1984, that the model was suitable for local or regional applications and that the term universal referred to the equation's general concept. For this reason, the USLE was revised several times after its publication, giving rise to modified versions, such as the Revised USLE (RUSLE), and alternative approaches, such as the Water Erosion Prediction Project (WEPP) (Laflen, 2003).

In recent years, the popularization of resourceful computers has allowed a wide dissemination of different modeling approaches. The systematic review of Borrelli et al. (2021) showed that 1,697 papers published 3,030 erosion modeling applications from 1994 to 2017. Their analysis showed the predominance of water erosion studies at a national scale and a clear temporal trend of increase in applications, ranging from 55 in 1994-1997 to 340 in 2014-2017. The authors also found as many as 435 different erosion models, with a clear

predominance of USLE-based approaches, corresponding to 41% of all applications. One of the reasons for the huge popularity of USLE-based methods is a relatively good tradeoff between input data requirements and the reliability of soil loss estimates. However, on the other hand, some limitations include: i) that applications often constitute extrapolations of the original farm-level scale for which the model was developed, and ii) the lack of a representation of sediment transport and deposition (Quine and van Oost, 2020). The second limitation is often addressed by including an external sediment routing scheme in the modeling framework, e.g. (Rompaey et al., 2001).

Historically, the development of erosion models was done in parallel to that of land surface models. Land surface models aim at integrating multiple equations to simulate terrestrial processes, including: the physical processes and chemical transformations (e.g., leaf growth), the column interactions (e.g., precipitation falling), the exchange between components (e.g., precipitated water entering the soil), the exchange of heat, and the motion of substances (e.g., water flows). Through these processes, calculations are constrained by fundamental physical laws, such as the conservation of mass and energy (Gettelman and Rood, 2016).. According to Doetterl et al. (2016), no land surface model currently contains a sufficiently detailed representation of the relationship between soil organic carbon and erosion, transport and deposition. The authors argue that this stems from inadequate input data, problems generalizing model parameters, an insufficient understanding of erosion processes, and large computational requirements.

Several works in the literature aimed at coupling erosion, transport and deposition to the biogeochemical cycle of carbon, with approaches varying in detail, complexity and scale. Lal (2003), for example, assumed a fixed delivery ratio of sediments to estimate aggregated global values of carbon displacement. This approach is similar to Chappell et al. (2015), which accounted for erosion as an extra carbon flux, but did not account for any particle transport. In the spatially explicit approach of Lugato et al. (2016), erosion was also accounted for as an extra flux, and particle transport to the rivers was assumed to be a constant share of the total eroded soil in each grid cell. Wang et al. (2017) also adopted a spatially explicit transport model, where the landscape was divided into regions named “virtual catchments” to accommodate an extra transport flux. At the continental scale, Borrelli et al. (2018a) coupled soil organic carbon data to an erosion model representing the transport of particles in a process-based distributed manner. This approach, however, is lumped and does not include any other elements of the carbon cycle. At the catchment scale, Nadeu et al. (2015) inserted

a spatially distributed routing scheme into a detailed carbon model in central Belgium. However, the limitation of this approach is the scale in which the model can be applied.

To address the problems mentioned above, Naipal et al. (2020) proposed the CE-DYNAM model for the non-Alpine part of the Rhine basin, a region whose area equals 185,000 km2. It combines erosion and transport modules to an emulator of the soil carbon dynamics of any land surface model based on CENTURY (Parton et al., 1983). CE-DYNAM uses a RUSLE-based approach adapted for modeling erosion at a large scale and a coarse spatial resolution (Naipal et al., 2015), and the transport module is a routing scheme that follows the topography to redistribute the soil particles (Naipal et al., 2016). These elements are formally incorporated into the model dynamics as additional fluxes between pools beyond those initially present in the first-order kinetics of CENTURY.

In its first version, CE-DYNAM was suitable for use at the catchment scale and could not be properly scaled to larger domains for several reasons. First, the run times ranged from hours to days at the catchment scale, indicating that calculations at larger domains could take weeks to complete. Such a limitation also precluded the execution of a calibration procedure since optimization routines often demand several function evaluations. Second, computer memory availability was a limiting factor since all datasets for the whole domain needed to be loaded and read before processing. Third, the code implementation of CE-DYNAM contained details that made it impractical for application when the number of sub-catchments in the domain increased. Finally, the intensity of the lateral fluxes had to be imposed and not calibrated with observed data, which could generate unrealistic results. All these problems were connected and indicated that, despite the potentialities, further work was necessary to improve CE-DYNAM.

1.3 . Policy initiatives for erosion and carbon

The New European Green Deal endorsed in 2020 includes the target to have all European soils healthy by 2050. Such an ambitious target contrasts with the current condition since 60-70% of the European soils are considered to be in an unhealthy state (Panagos et al., 2022b). The need for action in this field is therefore imperative. The European Union (EU) Soil Strategy for 2030 includes actions to facilitate the transition towards healthy soils through climate change mitigation, biodiversity preservation, soil restoration, and others. The Strategy, to-

gether with the recently established EU Soil Observatory, complements existing policies affecting soils at the European level, such as the Common Agricultural Policy (CAP) (Panagos et al., 2022b).

The Green Deal also extends the effort to recognize the importance of soils started by the EU in 2002, when the Thematic Strategy on Soil Protection, considered the first milestone for a continental soil protection policy, was established. Before this strategy, the funding for soil-related research came only indirectly through urban development, water protection, and biodiversity projects. A Soil Framework Directive was later proposed by the European Commission in 2005 after a Soil Advisory Forum with over 100 representatives and 350 experts produced final reports to guide the development of policies (Prokop, 2005). This directive faced significant opposition from countries such as France, Germany, the Netherlands, and others. Several points of disagreement existed, including doubts about the effectiveness of supranational policies for soils and the extra burden that it would put on countries that already had soil legislation at the national level. For these reasons, in 2014, the European Commission withdrew the proposal of the Soil Framework Directive (Azam, 2016).

The ecosystem services provided by soils are relevant to 13 out of the 17 Sustainable Development Goals (Keesstra et al., 2016), and in the EU, they are essential to achieve a variety of actions listed in the Green Deal, such as the achievement of climate neutrality by 2050 envisaged by the Climate Law. Soils are also explicitly mentioned in the Biodiversity Strategy, the Farm to Fork Strategy, and the Zero Pollution action plan. For example, the Farm to Fork Strategy aims to improve food systems by making them more fair and healthy. It proposes reducing fertilizer usage by 20% and halving nutrient losses and the use of pesticides by 2030. In the Biodiversity Strategy, the reversion of negative trends will be addressed by aiming at land degradation neutrality, reducing soil erosion, increasing soil organic matter, and restoring severely eroded agricultural land (Montanarella and Panagos, 2021).

Another central policy to affect agricultural land management in the EU is the CAP. Payments are currently linked with Good Agricultural and Environmental Conditions (GAECs), which establish minimum land management conditions for crop systems (Panagos et al., 2022b). One example of a GAEC introduced in the upcoming CAP 2023-2027 is the adoption of cover crops (CCs) to protect soils during rainy periods. The implementation was supported by a scenario analysis from the European Commission, which indicated that erosion rates in arable lands and permanent crops could reduce by 15 and 30%, respectively, with the adoption of CCs at a cover rate of 75% (Panagos et al., 2021). Consequently, even though the new

CAP rules can reduce the effective CC area in some farms, the net area of CCs across the EU is expected to increase after the updates.

Besides, the European climate target also intersects with soil, as sustainable management practices can increase the organic carbon stocks in agricultural land. To this end, policies related to the Green Deal are dependent on scientific evidence about alternative sustainable management strategies to provide environmental benefits. Actions focused on carbon, for example, include developing soil management plans by Member States, planting 3 billion trees across the European Union, promoting wetlands maintenance, and reducing CO_2 emissions. In this context, advancing modeling studies to represent management alternatives can be helpful to inform the decisions to be made, and to understand the limits of proposed actions.

1.4 . Thesis structure

The present PhD thesis had three objectives. The first objective (**Chapter 2**) was to propose an alternative formulation for CE-DYNAM to allow the upscaling of its calculations to the continental level. This was done by formalizing the model formulation, representing the model's first-order dynamics in matrix form, and introducing a calibration procedure based on sediment discharge measurements. In this chapter, a second version of CE-DYNAM (i.e., v2) was presented, which consists of v1 plus the matrix form and a few modifications. The first chapter dealt with *mechanistic modeling*.

The second objective (**Chapter 3**) was to generate the first satellite-based map of cover crops in Europe. Since the adoption of cover crops is a management action with direct impacts on both the soil organic carbon cycle and in erosion prevention, this chapter aimed at gaining more understanding of their spatial distribution at a continental scale. The second chapter dealt with *empirical modeling*.

Finally, the third objective (**Chapter 4**) was to evaluate the potential impacts of cover crops on soil carbon stocks and the export of particulate organic carbon to the oceans in Europe. This work combined the developments of Chapter 1 and Chapter 2, and a third version of CE-DYNAM (i.e., v3) was presented. This version consisted of v2 plus several major modifications added to improve the model's physical representation of erosion, transport, and deposition processes. The third chapter dealt with *the evaluation of policy alternatives*.

2 - Matrix representation of lateral soil movements: scaling and calibrating CE-DYNAM (v2) at a continental level

This chapter was written by Arthur Nicolaus Fendrich, with contributions from Philippe Ciais, Emanuele Lugato, Marco Carozzi, Bertrand Guenet, Pasquale Borrelli, Victoria Naipal, Matthew McGrath, Philippe Martin, and Panos Panagos.

Promoting sustainable soil management is a possible option for achieving net-zero greenhouse gas emissions in the future. Several efforts in this area exist, and the application of spatially explicit models to anticipate the effect of possible actions on soils at a regional scale is widespread. Currently, models can simulate the impacts of changes on land cover, land management, and the climate on the soil carbon stocks. However, existing modeling tools do not incorporate the lateral transport and deposition of soil material, carbon, and nutrients caused by soil erosion. The absence of these fluxes may lead to an oversimplified representation of the processes, which hinders, for example, a further understanding of how erosion has been affecting the soil carbon pools and nutrient through time. The sediment transport during deposition and the sediment loss to rivers create dependence among the simulation units, forming a cumulative effect through the territory. If, on the one hand, such a characteristic implies that calculations must be made for large geographic areas corresponding to hydrological units, on the other hand, it also can make models computationally expensive, given that erosion and redeposition processes must be modeled at high resolution and over long time scales. In this sense, the present work has a three-fold objective. First, we provide the development details to represent in matrix form a spatially explicit process-based model coupling sediment, carbon, and erosion, transport and deposition processes (ETD) of soil material in hillslopes and valley bottoms (i.e., the CE-DYNAM model). Second, we illustrate how the model can be calibrated and validated for Europe, where high-resolution datasets of the factors affecting erosion are available. Third, we presented the results for a depositional site, which is highly affected by incoming lateral fluxes from upstream lands. Our results showed that the benefits brought by the matrix approach to CE-DYNAM enabled the before precluded possibility of applying it to a continental scale. The calibration and validation procedures indicated: i) a close match between the erosion rates calculated and previous works

in the literature at local and national scales; ii) the physical consistency of the parameters obtained from the data; and iii) the capacity of the model in predicting sediment discharge to rivers in locations observed and unobserved during its calibration (Model efficiency (ME) $= 0.603$, $R^2 = 0.666$; and ME $= 0.152$, $R^2 = 0.438$, respectively). The prediction of the carbon dynamics on a depositional site illustrated the model's ability to simulate the non-linear impact of ETD fluxes on the carbon cycle. We expect that our work advances ETD models' description and facilitates its reproduction and incorporation in land surface models such as ORCHIDEE. We also hope that the patterns obtained in this work can guide future ETD models at a European scale.

2.1 . Introduction

The adoption of more sustainable land management actions constitutes a critical alternative for mitigating climate change and sustaining food production (Roe et al., 2019). Soils constitute a vital carbon (C) pool for the world, storing 1500-2400 petagrams of carbon (PgC), more than the atmosphere (589 PgC) the surface ocean (900 PgC) together, and the way humans interact with soils affects how soils and the atmosphere interact, including the sequestration of carbon (Ciais et al., 2013). It is understood that even minor disturbances to soil pools can have significant impacts on the global C cycle: increases of 4‰ in global agricultural stocks, for example, could result in additional C sequestration of 2 to 3 PgC per year, which would contribute significantly to the Paris agreement targets (Guenet et al., 2020; Minasny et al., 2017; Soussana et al., 2019). Alternatives to rapidly increase the content of soil organic matter, and consequently the C sequestration from the atmosphere, include conservation agriculture (e.g., the retention of residues and zero or no-tillage) (Robert, 2001), agroforestry and afforestation. In the future, however, the projected population growth poses an increasing demand for food, feed, energy, and water, resulting in additional pressures that, if not properly dealt with, can even aggravate the problem (IPCC, 2019). An iconic example is the Southeastern Amazon forest, which has long been understood as a C sink, but after decades of deforestation, it is becoming a source of C to the atmosphere (Gatti et al., 2021; Nobre et al., 2016).

One of the possible strategies for evaluating the impacts of different alternatives on the C stocks is the use of numerical models that represent the physical, chemical, and biological processes of the soil-plant-atmosphere system, such as fixation by plants for biomass

growth and the respiration by microorganisms (Gettelman and Rood, 2016). Models representing the interaction between soils and the atmospheric system allow the evaluation of how future climate change will impact soils and the opposite relationship. For example, land surface models (LSM) have allowed studies on different topics, such as assessing the impacts of climate change on crops, habitat, and water availability (Leng and Hall, 2019; Schewe et al., 2019; Hamaoui-Laguel et al., 2015; Bonan and Doney, 2018), evaluating strategies to achieve global environmental targets (Harper et al., 2018; Chang et al., 2021), forecasting future scenarios of change (Friedlingstein et al., 2006; Friedlingstein, 2015), among others. However, the implementation state of LSMs currently does not cover some relevant processes, such as lateral displacement of nutrients in the soil due to erosion, transport, and deposition (ETD) processes (Quine and van Oost, 2020). ETD is argued to affect the carbon cycle dynamically during its occurrence by inducing lateral fluxes of C in the landscape and vertical fluxes between soil layers (Lal, 2003; Lugato et al., 2018), and their absence in LSMs leads to an oversimplified representation of the reality. The modeling complexity, along with the scarcity of empirical data for the phenomenon and the non-standardized nomenclature in the literature, hinder, for example, a further understanding of how erosion has been affecting the soil C pools through time (Lal, 2019; Lugato et al., 2018; Wang et al., 2017; van Oost et al., 2007).

Including the complex ETD-related processes into existing LSMs comes at the cost of increasing the inherent technical complexity of these mechanistic models, such as requiring massive amounts of codes, demanding costly computational resources, and being hard to diagnose thoroughly (Lu et al., 2020). For example, even without ETD-related processes, existing LSMs are so complex with their detailed soil-vegetation-atmosphere feedbacks and multitude of spatial or temporal scales that simulations often must be performed repeatedly for hundreds or thousands of years until a stable condition is reached (Huang et al., 2018). In these cases, calculations can take hundreds of processor hours, and researchers often adopt less detailed or simplified processes to avoid prohibitively slow simulation times (Washington et al., 2008). Practically, such technical problems may hinder their operation by users, which are often individuals with different backgrounds and abilities. Since such problems can have an impact on model testing, validation, and ultimately acceptance by the scientific community, approaches to overcome them have been studied in the recent past. It is, for example, the case of the matrix approach, which consists of representing all carbon fluxes explicitly in matrix form (Luo et al., 2017), which has been reported to increase modularity, facilitate diagnostics, and accelerate spin-up calculations (Lu et al., 2020; Xia et al., 2012; Huang et al., 2017, 2018). For ETD-related processes, the incorporation and development of

such approaches are advisable, encouraged and necessary to enable the complexity of representing the vertical and lateral dynamics of C and sediments on the landscape.

In this work, we address the problem of scaling the calculations of CE-DYNAM, a hybrid empirical-mechanistic ETD model based on a physical emulator of the carbon cycle in soils (Naipal et al., 2020) to a continental scale. First, we describe the model formulation and show how the matrix approach leads to a sparse linear system, thus making calculations feasible. We expect our mathematical development of CE-DYNAM to facilitate its reproduction and incorporation in LSMs such as ORCHIDEE, DayCent, and others. We then calibrated the model for the study area (i.e., Europe) for the last 150 years using climate forcings with a monthly temporal and a 0.125° spatial resolution (approx. 12.5km at the Equator) using sediment concentration in rivers data collected on the field. Comparing the predictions against such observed values is important to evaluate the cumulative effect of all model assumptions, as well as its performance on catchments with different characteristics. Internal and external validation of the results is presented to show their consistency and physical realism. Finally, we exemplify the practical use of CE-DYNAM by presenting the results of the impact of ETD-related processes in a chosen depositional area in the territory. With the model in a matrix form, the calibration, and the pattern obtained at the depositional area, we expect to form the basis for future large-scale model applications.

2.2 . Materials and Methods

2.2.1 . Methodological proposal: the matrix approach

Definitions

The CE-DYNAM model (Naipal et al., 2020) consists of coupling erosion and transport modules to the soil carbon dynamics of any land surface model based on CENTURY (Parton et al., 1983, 1988). Typically, CE-DYNAM uses the Revised Universal Soil Loss Equation (RUSLE - Renard, 1997), an approach adapted for predicting erosion at a large scale and a coarse spatial resolution, but any other existing option such as the LISEM model (de Roo et al., 1998) could be used. The transport module is a topography-based routing scheme, which uses the altitude (or an approximate digital elevation model) to distribute the sediments and their corresponding organic carbon. The scheme is calibrated with field sediment discharge data to generate realistic values, and the elements are incorporated into the soil organic C dynamics

as additional fluxes between pools beyond those initially present in the first-order kinetics of CENTURY (Naipal et al., 2015, 2016). As a remark, CE-DYNAM could be coupled to other carbon models as long as they adopt a first-order kinetics. Some advantages of the coupled approach of CE-DYNAM include the current incorporation of interactions such as the feedback between land use, climate, and erosion (Borrelli et al., 2020; Quine and van Oost, 2020), and the potential for the future implementation of other components such as soil properties. Figure 2.1 presents a simplified representation of all fluxes of hillslopes and valley bottom soil pools in CE-DYNAM.

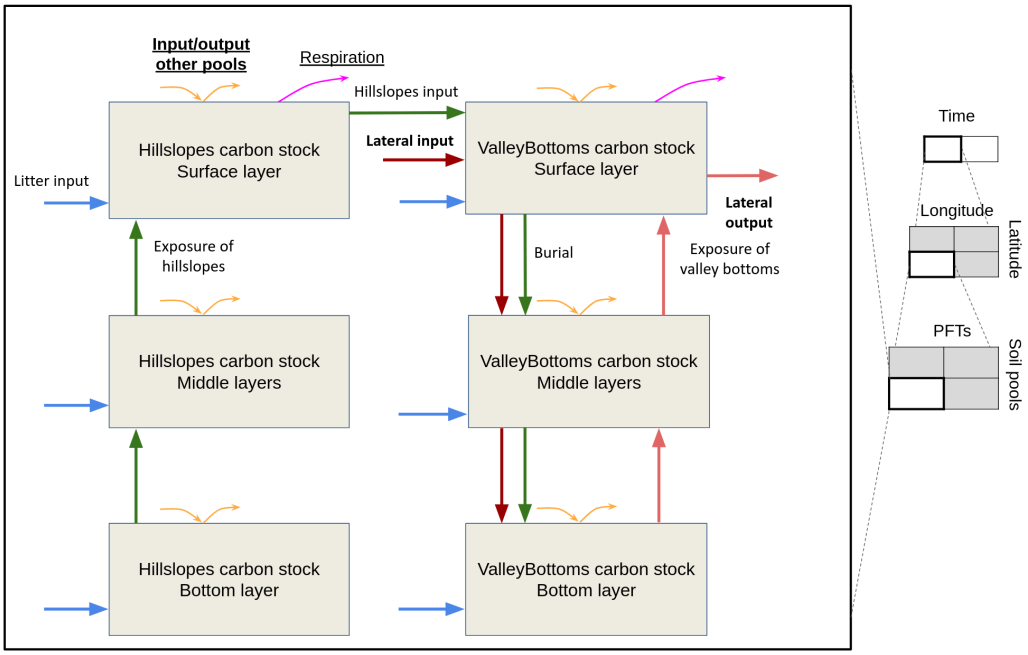


Figure 2.1: A simplified representation of all fluxes in CE-DYNAM. All colors represent the same flux (e.g., the blue arrow represents the input from litter). The example shown corresponds to a specific moment in time, spatial location, plant functional type (PFT), and soil pool. All fluxes with written descriptions are directly affected by the parameters to be calibrated (i.e., γ_1, γ_2 etc.), except for those with an underline. Fluxes whose description is in **bold** interact with one or more spatial locations, soil pools, or PFTs (right, gray squares)

CE-DYNAM has only been applied at a local scale, such as in the non-alpine region of the Rhine basin (whose total area equals 185.000 km^2) (Naipal et al., 2020). However, scaling the model to the continental scale, where the area can be tens of times larger than the application mentioned above, still faces practical implementation difficulties that we address in our

current work. A careful evaluation of CE-DYNAM's original implementation allows the identification of three important remarks. First, from a computational point of view, the original implementation requires the storage of a large amount of data in the computer's memory for its execution, which in practice becomes prohibitive as the geographical area or spatial resolution increases. Second, the original strategy is based on the subdivision of the problem in smaller and adjacent units – generally the sub-basins of a hydrographic basin. This procedure naturally restricts the ability to distribute the solution over different processing units and requires the continuous execution of additional steps of integration of all smaller units, which leads to a significant performance reduction. Third, the equilibrium calculation procedure of the original method consists of the successive iteration of the model, which can be very inefficient (Huang et al., 2018). Alternatives for these problems are more easily perceived when the models' mathematical notation is properly developed and stated.

Thus, to solve the problems above, we first clarify the notation to facilitate the comprehension of details and ensure reproducibility. We do so by adopting a general description and, when necessary, including examples based on ORCHIDEE (Krinner et al., 2005) to illustrate the concepts. The formulation accompanies Table 2.1, containing all input variables for the model, which helps clarification. The indices in Table 2.1 refer to five dimensions: the soil pool (c), the spatial location (x, interpreted here as a point of a lattice X representing an area in the surface), the plant functional type (PFT) (p), the soil depth (d). Besides those dimensions, most variables also evolve in time (t). Some data sets are assumed constant on one or more dimensions during simulations: the geographic area of each cell, for example, varies in space but does not change according to the soil pools, PFTs, soil depth, nor time.

Table 2.1: The external input variables for CE-DYNAM calculation.

Notation	Description	Source
$u_{x,p_l}(t)$	The percentage of each plant functional type (PFT) in each cell. $0 < u_{x,p_l}(t) < 1$, so one cell cannot be more than 100% covered by a PFT; and $\sum_l u_{x,p_l}(t) = 1$.	This information comes from the land surface model.
a_x	The geographic area (km^2) of each cell.	Derived from each cell's bounding box.
b_{x,d_k}	The bulk density (g/cm^3).	Several potential sources are available, for example the Global Soil Dataset for Earth System Modeling (≈1km) (Shangguan et al., 2014) and SoilGrids (250m) (Poggio et al., 2021).
α_x	The depth to bedrock (cm).	SoilGrids (250m) (Poggio et al., 2021).
ν_{x,d_k}	The soil organic carbon stock (tonnes).	Derived from SoilGrids (250m) (Poggio et al., 2021).
$\kappa^{i,j}_{x,p_l}(t)$	When $i = j$, it represents an output rate from the pool i. When $i \neq j$, it represents a transfer rate between (from) carbon pool c_i and (to) carbon pool c_j (1/day).	These rates are calculated from the output of the land surface model.
$\rho^i_{x,p_l}(t)$	The respiration rates of carbon pool c_i (1/day).	Idem as above.
$I^j_{x,p_l}(t)$	The input from litter pools to carbon pool c_j (gC/(m^2.day)).	Idem as above.
$e_{x,p_l}(t)$	The average erosion rate (tonnes/(ha.day)).	Calculated from any erosion model such as the Universal Soil Loss Equation, the Water Erosion Prediction Project etc.
s_x	The terrain slope (degrees).	Derived from any digital elevation model, such as from the Shuttle Radar Topography Mission (Farr et al., 2007).
$1/w_x$	The adimensional flow accumulation (i.e., the cumulative number of upstream drainage cells).	Idem as above, preferably from the same source as the terrain dataset.
h_x	The fraction of each cell belonging to hillslopes. $(1 - h_x)$ is the fraction of valley bottoms.	Pelletier et al. (2016).
l_x	The river width (m).	Derived from HydroSheds (Lehner et al., 2008).

Mimicking the carbon dynamics of the LSM (in our case, ORCHIDEE) is the most important pillar of CE-DYNAM (Naipal et al., 2020). In general, we can represent the soil carbon pools setting of such LSMs with a set $\mathrm{C}_s = \{c_1, c_2, \ldots, c_n\}$. However, in comparison to the LSM in which it is based, CE-DYNAM makes additional assumptions to those described above. One of these assumptions is that the soil carbon pools are divided into two fractions: hillslopes and valley bottoms (i.e., $\mathrm{C}_s = \mathrm{C}_h \bigcup \mathrm{C}_v$), in such a way that the original number of soil carbon pools is twice the number of that in the LSMs. Such an assumption affects the original calculation depending on the fraction under consideration. For the hillslopes, calculations are modified by the inclusion of an extra flux proportional to the erosion predicted by a chosen model such as the RUSLE. For the valley bottoms, such a flux from hillslopes becomes a new input, and a new lateral dynamics of sediments across the landscape induced by the terrain slope (s_x) and the flow accumulation ($1/w_x$), is added. These lateral dynamics give rise to most computational challenges in CE-DYNAM since they make the stock in one simulation unit dependent on its neighbors.

Another assumption introduced by CE-DYNAM is a discretization of soil depth, which allows the evaluation of the vertical movement of carbon in layers even when the LSM does not. This is done by first setting $m = 3$ soil layers (that is, surface, middle and bottom layer) and then defining a set $\mathrm{D} = \{d_1, d_2, \cdots, d_m\}$ of soil layers. For example, if one let $d_1 = 10cm$, $d_2 = 20cm$ and $d_3 = 30cm$, then one is identifying the segment of soil from zero to $10cm$ as the surface, the segment from $10cm$ to $30cm$ as the middle layer and from $30cm$ to $60cm$ as the bottom layer. Throughout this text, we also use the symbol d_k, $k \in \{1, 2, \cdots, m\}$, to refer to the k-th layer. Then, the input from litter to soil pools is distributed along D to calculate the share of input to each soil layer. We describe such a vertical discretization procedure in Equation 2.2.1, and we denote the vertically-discretized version of $I^j_{x,p_l}(t)$ as $I^{*j}_{x,p_l,d}(t)$ (Equation 2.5).

Because the LSM used in this work is based on CENTURY, carbon pool kinetics will always follow a first-order differential equation. Furthermore, soil carbon is divided into three pools (active, slow, passive) with different turnover rates that vary with temperature, moisture, clay content, and other modifiers (e.g., tillage) (Camino-Serrano et al., 2018). The set of $v = 15$ plant functional types used to represent land cover in the model is denoted here as $P = \{p_1, p_2, \cdots, p_v\}$. Then, for a fixed layer $d_k \in D$, a fixed lattice point $x \in X$, a fixed PFT $p_l \in P$, a fixed pool $c_i \in C_s$ and a fixed time t we let $S^i_{x,p_l,d_k}(t)$ denote its carbon stock.

The formulas for the CE-DYNAM rates are detailed in later sections of the text. However,

we can essentially represent how the model evolves in time with Equation 2.1. While such a representation omits most model dimensions, it's useful to clarify its dynamics as that of a linear and non-autonomous model (Sierra et al., 2018, Table 1). As we will describe in the following subsections, the coupling of erosion-related processes will always respect this general structure, with the changes consisting of modifications to each of its elements according to the particular case.

$$\frac{dS}{dt} = \mathrm{I}(t) - \kappa(t) \cdot S(t) \tag{2.1}$$

with $S(t)$ denoting the carbon stock at time t in the pool, $\mathrm{I}(t)$ denoting all the pool's input, and $\kappa(t)$ denoting the output rates. In the equilibrium calculation, the model was iterated several times over the period 1860-1869 until convergence to the pullback attracting trajectory (Sierra et al., 2018). In the transient period, all the elements on the right-part of the equation will be known and dS/dt calculated will correspond to the increment in carbon stocks at each time step. Essentially, we are interested in evaluating how the carbon stocks $S(t)$ change over the transient period. Through the rest of the text, we frequently refer to Equation 2.1 as the basis to form the carbon budget in all cases.

In the matrix approach, we discretize Equation 2.1 and represent all fluxes between pools as a linear system. Hypothetically, if no fluxes between pools existed, we would have

$$S(t + \Delta t) = S(t) + I(t) \cdot \Delta t - A(t)S(t) \cdot \Delta t, \tag{2.2}$$

where $S(t) = \left[S^{i}_{x,p_l,d_k}(t)\right]_{\substack{x\in X \\ l\in[v] \\ k\in[m] \\ i\in[n]}} \in \mathbb{R}^{|X|\times v\times m\times n}$, $I(t) = \left[I^{i}_{x,p_l,d_k}(t)\right]_{\substack{x\in X \\ l\in[v] \\ k\in[m] \\ i\in[n]}} \in \mathbb{R}^{|X|\times v\times m\times n}$ and $A(t)$ is a diagonal matrix with diagonal $\left[\kappa^{i,i}_{x,p_l,d_k}(t)\right]_{\substack{x\in X \\ l\in[v] \\ k\in[m] \\ i\in[n]}} \in \mathbb{R}^{|X|\times v\times m\times n}$.

However, interactions tend to be complex in more general situations. The following sections show that the routing scheme for valley bottoms creates a dependence between pools of different grid cells, PFTs, and soil layers. While such a property replaces several off-diagonal zero elements of $A(t)$ by non-zero rates, it still preserves the inherently sparse structure of $A(t)$.

Next, we detail how the elements of $A(t)$ and $I(t)$ can be calculated. For simplicity, we exemplify with the first timestep of the equilibrium period ($t = t_0$), but calculations are anal-

ogous for all timesteps.

Vertical discretization

As mentioned in subsection 2.2.1, CE-DYNAM vertically discretizes the soil, which has a direct impact on the respiration, erosion, and turnover rates of the original LSM. An exponential increase of the profile depth is assumed, so each soil layer thickness in the discretization profile is calculated from the depth to bedrock, α_x, using two real-valued parameters γ_1 and r:

$$d_{m-k+1} = \int_{z=(k-1)/m}^{z=k/m} \alpha_x \cdot \exp(\gamma_1 + r \cdot z)\mathrm{d}z = \frac{\alpha_x}{r} \cdot \left[\exp\left(\gamma_1 + r \cdot \frac{k}{m}\right) - \exp\left(\gamma_1 + r \cdot \frac{k-1}{m}\right)\right] \tag{2.3}$$

$$\forall k = 1, 2, \cdots, m.$$

For any choice of γ_1, the parameter r is calculated by constraining the sum of all vertical layers to match the total distance to bedrock. By using the general properties of definite integrals, it is possible to show that it can be analytically calculated with the closed-form solution

$$r = -\exp(\gamma_1) - \mathrm{W}\left[-\exp(\gamma_1 - \exp(\gamma_1))\right], \tag{2.4}$$

where $\mathrm{W}(\cdot)$ representing the Lambert W function (see Corless et al., 1996). The example notation of Equation 2.3 and Equation 2.4 shows an important property of the vertical discretization scheme: the γ_1 parameter depends solely on the soil discretization setting, which is assumed to be identical for all cells in CE-DYNAM. For this reason, there is a single γ_1 value independent of all factors being different (e.g., the spatial location, the PFT or the variable to be discretized). Besides, this property also means that the vertical discretization is not scale-invariant, and thus the depth scheme must be defined with extra care. In Figure 2.2, we show how a possible vertical profile of depth equal to 2m varies with γ_1: for values closer to zero (left), the profile tends towards a flat one, while for larger values (right), the model tends to calculate a smaller surface layer. The realistic choice of γ_1 must come from the model calibration procedure.

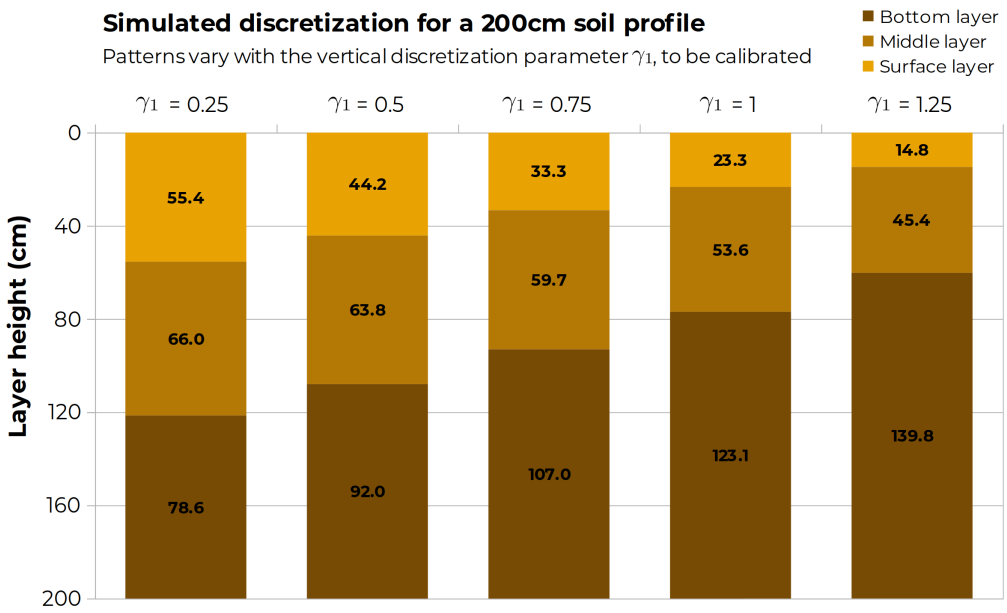


Figure 2.2: Possible options for the vertical discretization parameter in CE-DYNAM. The vertical axis shows the layer heights in centimeters, and the horizontal axis shows some possible γ_1 values.

In the carbon simulation, the input from litter to soil pools is also vertically discretized. This is done by multiplying the original quantity (i.e., $I^j_{x,p_l}(t)$) by the percentage of soil organic carbon in each soil layer (Poggio et al., 2021).

$$I^{*j}_{x,p_l,d_k}(t) = I^j_{x,p_l}(t) \cdot \frac{\nu_{x,d_k}}{\sum_d \nu_{x,d}}, \tag{2.5}$$

The erosion rates correspond to the share of the total mass of soil that is eroded in each vertical layer. Since not all the carbon eroded in hillslopes goes to valley bottoms, the term is multiplied by a fraction from zero to one, assumed to vary with terrain slope and land cover. A different curve is assumed for forests, croplands, and grasslands, and their calibration is

made using field observations.

$$\lambda_{x,p_l,d_k}(t) = \overbrace{\frac{e_{x,p_l}(t)}{\underbrace{b_{x,d_k} \cdot d_k \cdot h_{x,p_l} \cdot a_x}_{\text{Total mass of soil}}}}^{\text{RUSLE rate}} \cdot \overbrace{\frac{1}{1+\exp(g_f(s_x))}}^{\text{\% of erosion that goes to valley bottoms}}, \tag{2.6}$$

with the multiplication by h_{x,p_l} varying for hillslopes and valley bottoms according to their fractions, and $g_f(s_x)$ being the weighted sum of the different smoothing function for forests, croplands, and grasslands multiplied by their corresponding land cover fractions. Although not explicit in the notation, such a function also varies on time, since land cover varies each year.

Fluxes: hillslopes soil carbon pools

Bottom soil layer: d_m We describe the carbon dynamics in hillslopes in terms of three general pools $c_1, c_2, c_3 \in \mathrm{C}_h$, which can be interpreted in terms of the active, slow, and passive soil pools of ORCHIDEE. For the deepest soil layer, the rearrangement of Equation 2.1 leads to the following equations for c_1, c_2 and c_3, respectively:

$$\frac{dS^1_{x,p_l,d_m}(t_0)}{dt} = \left(\overbrace{\sum_{c_i \in C_l} I^{*1}_{x,p_l,d_m}(t_0)}^{\text{Input: litter pools}} + \overbrace{\kappa^{2,1}_{x,p_l}(t_0) \cdot S^2_{x,p_l,d_m}(t_0)}^{\text{Input: } c_2 \text{ pool}} + \overbrace{\kappa^{3,1}_{x,p_l}(t_0) \cdot S^3_{x,p_l,d_m}(t_0)}^{\text{Input: } c_3 \text{ pool}} \right) - \left(\overbrace{\rho_{x,p_l}(t_0)^1}^{\text{Respiration rate}} + \overbrace{\kappa^{1,2}_{x,p_l}(t_0)}^{\text{Output: } c_2 \text{ pool}} + \overbrace{\kappa^{1,3}_{x,p_l}(t_0)}^{\text{Output: } c_3 \text{ pool}} + \overbrace{\lambda_{x,p_l,d_m}(t_0)}^{\text{Output: erosion b} \rightarrow \text{t}} \right) \cdot \overbrace{S^1_{x,p_l,d_m}}^{c_1 \text{ stock}} \tag{2.7}$$

$$\frac{dS^2_{x,p_l,d_m}(t_0)}{dt} = \left(\sum_{c_i \in C_l} I^{*2}_{x,p_l,d_m}(t_0) + \kappa^{1,2}_{x,p_l}(t_0) \cdot S^1_{x,p_l,d_m} + \kappa^{3,2}_{x,p_l}(t_0) \cdot S^3_{x,p_l,d_m}(t_0) \right) - \left(\rho^2_{x,p_l}(t_0) + \kappa^{2,1}_{x,p_l}(t_0) + \kappa^{2,3}_{x,p_l}(t_0) + \lambda_{x,p_l,d_m}(t_0) \right) \cdot S^2_{x,p_l,d_m}(t_0) \tag{2.8}$$

$$\frac{dS^3_{x,p_l,d_m}(t_0)}{dt} = \left(\sum_{c_i \in C_l} I^{*3}_{x,p_l,d_m}(t_0) + \kappa^{1,3}_{x,p_l}(t_0) \cdot S^1_{x,p_l,d_m} + \kappa^{2,3}_{x,p_l}(t_0) \cdot S^2_{x,p_l,d_m}(t_0) \right) - \left(\rho^3_{x,p_l}(t_0) + \kappa^{3,1}_{x,p_l}(t_0) + \kappa^{3,2}_{x,p_l}(t_0) + \lambda_{x,p_l,d_m}(t_0) \right) \cdot S^3_{x,p_l,d_m}(t_0) \tag{2.9}$$

Since CE-DYNAM does not affect litter pools, all quantities on the equations above should be known, except the three hillslope soil carbon pools in equilibrium calculation. In Equation 2.7, we denote as b $\rightarrow$ t the flux from the bottom layer to the layer above.

Middle and top soil layers In hillslopes, the structure for middle and top soil layers will be identical as for bottom layers, except for the b $\rightarrow$ t loss from the layers below (i.e., the fourth term of Equation 2.7), which becomes a new input to the layers above. This results in an additional input equal to $\lambda_{x,p_l,d_{k+1}}(t_0) \cdot S^j_{x,p_l,d_{k+1}}(t_0)$, in case of pool c_j and depth d_k. One important final remark is that, for the top soil layer, the interpretation of the "Output: erosion b $\rightarrow$ t" rate becomes "Output: erosion hillslopes $\rightarrow$ valley bottoms".

Fluxes: valley bottoms soil carbon pools

Preliminary assumptions In the hillslopes soil carbon pools described above, the movement of C was spatially static, which means that all calculations were performed within the same spatial unit (i.e., grid cell). However, the physical definition of valley bottoms extends

the movement to other cells since connected areas exchange sediments and C according to the terrain and land cover configuration. This characteristic is incorporated into the CE-DYNAM model by defining a routing scheme that transports sediments along the landscape.

To represent the lateral transport, a new rate τ_x derived from the sediment residence time is added. Its calculation is performed as:

$$\overbrace{\tau_x}^{\text{Sediment rate}} = \underbrace{\frac{1}{g_\tau(1/w_x)}}_{\text{Residence time}} \tag{2.10}$$

with $g_\tau(1/w_x)$ representing a smoothing function between the sediment residence time and the flow accumulation (i.e., upstream area), to be calibrated from the observations. In this work, we adopted 3rd degree B-Spline to represent all smoothing functions. Besides, the flow accumulation information is also used in the routing scheme to generate an approximated digital elevation model, w_x. Such an approximation is used instead of the original terrain to ensure a hydrologically consistent topography for the lateral movements.

Also, let $\mathrm{P}_x^+(t)$ be the number of non-zero PFTs in cell x at time t and $\mathrm{Q}(x)$ be the set of queen neighbors (Figure 2.3) (Quinn et al., 1991) of a given cell x formed as:

$$\mathrm{Q}(x) = \{y : y - x \in \{-1, 0, 1\} \times \{-1, 0, 1\}, y \neq x\}$$

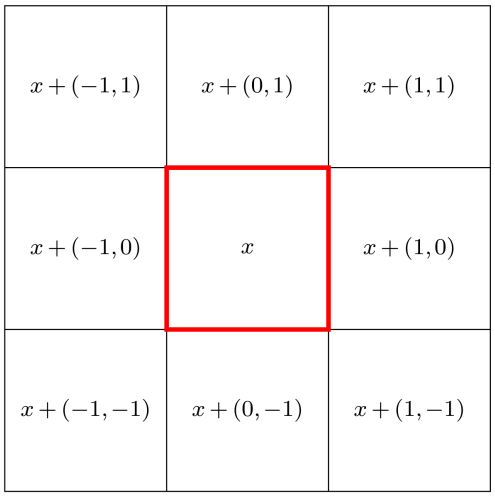


Figure 2.3: Queen neighbor setting used for the routing scheme. The center cell (red) is x.

With these definitions, the routing scheme for a given cell consists of two elements incorporated into its C balance. First, at the surface depth, d_0, there is loss from the cell to its neighbors, but some definitions are necessary to dictate how the process occurs:

1. The routing scheme works only within the same carbon pool. For example, the *active carbon* routed from one cell is added exclusively to the *active carbon pool* of its neighbor cells.
2. The C from one PFT in the source cell is transferred equally to all non-zero PFTs of the target cell.
3. The bare soil PFT (conventionally denoted here by p_0) loses and gains no C on the routing scheme.

The rate of routed C from a PFT p_r of the source cell x to a PFT p_l of the target cell y can then be calculated as:

$$\overbrace{\theta_{[\ \underbrace{x\ ,\ y}_{\text{Source, Target}}\ ]}}^{\text{Lateral transfer rate}} = \overbrace{\mathbb{1}_{(w_y<w_x)}}^{\text{Indicator function}} \cdot \overbrace{\frac{w_x - w_y}{||x-y||_2}}^{\text{Slope: rise/run}} \tag{2.11}$$

$$\overbrace{\zeta_{x,p_r,y,p_l}(t)}^{\text{Rate of routed carbon}} = \overbrace{\mathbb{1}_{(p_l \neq p_0, p_r \neq p_0)}}^{\text{Indicator}} \cdot \overbrace{\tau_x}^{\text{Sediment rate}} \cdot \overbrace{\frac{\theta_{x,y}}{\sum_{y \in Q(x)} \theta_{x,y}}}^{\text{Local \% of lateral transfer}} \cdot \overbrace{\frac{u_{y,p_l}(t)}{\sum_{s=1}^{v} u_{y,p_s}(t)}}^{\text{Local \% of target PFT}} \tag{2.12}$$

where the indicator function, $\mathbb{1}$, equals to one when the condition is met or zero otherwise. Together, Equation 2.11 and Equation 2.12 result in the important remark: the total loss of C in PFT $p_r \neq p_0$ of the source cell is equal to τ_x times the corresponding C stock at the surface, which varies according to the pool under consideration (for example, $S^1_{x,p_r,d_0}(t_0)$ for pool c_1). Also, the flux is equal to zero for the remaining case of $p_r = p_0$ since the bare soil does not participate in the routing scheme. At the surface, the equilibrium value of C stock in one cell and PFT will depend on the equilibrium value of C stock in all the PFTs of all its neighbors. This property of the routing scheme is essential and makes several zero off-diagonal elements be represented as rates from/to different grid cells, PFTs or soil layers.

Besides, despite affecting more directly the soil surface layer, the routing scheme is also assumed to affect the vertical movement of C earlier described in Equation 2.2.1 for the case of hillslopes. The same total rate routed from one PFT to the neighbors also moves through the layers, from the bottom to the top (b $\rightarrow$ t) (i.e., subsoil exposure). In the other way (t $\rightarrow$ b), the rate received from the neighbors is transmitted vertically from each layer to the layer below (i.e., burial). The only exception is naturally d_m, which has no layers below. Such input rate to p_l can be denoted as: $\sum_{y \in Q(x)} \sum_{r=1}^{v} \zeta_{y,p_r,x,p_l}(t)$.

Top soil layer: d_0 The equations for the C dynamics in valley bottoms can be obtained by putting the new fluxes along with the other ones from the original LSM. This implicitly assumes that litter input and PFT in valley bottoms are the same as in the standard LSM. Again, we make this section using a general notation of $c_1, c_2, c_3 \in \mathrm{C}_h$ and its respectively correspondent pools $c_4, c_5, c_6 \in \mathrm{C}_v$. For example, if c_1 is the hillslope soil active carbon pool, then c_4 is the valley bottoms soil active carbon pool. Below, we describe the fluxes for PFT p_l of pool c_4 using element-wise notation. For the topsoil layer, d_0, we have input from below but not from above, and also have inputs from some neighbor cells via the routing scheme and losses for other neighbors for the same reason.

$$\frac{dS^4_{x,p_l,d_0}(t_0)}{dt} = \Bigg(\overbrace{\sum_{c_i \in C_l} I^{*4}_{x,p_l,d_0}(t_0)}^{\text{Input: litter pools}} + \overbrace{\kappa^{6,4}_{x,p_l}(t_0) \cdot S^6_{x,p_l,d_0}(t_0)}^{\text{Input: } c_6 \text{ pool}} + \overbrace{\kappa^{5,4}_{x,p_l}(t_0) \cdot S^5_{x,p_l,d_0}(t_0)}^{\text{Input: } c_5 \text{ pool}}$$
$$+ \overbrace{\lambda_{x,p_l,d_0}(t_0) \cdot S^1_{x,p_l,d_0}(t_0)}^{\text{Input: hillslopes}} + \overbrace{\sum_{y \in Q(x)} P^+_y(t_0) \cdot \tau_x \cdot S^4_{x,p_l,d_1}(t_0)}^{\text{Input: vertical flow b} \rightarrow \text{t}}$$
$$+ \overbrace{\sum_{y \in Q(x)} \sum\nolimits_{r=1}^{v} \zeta_{y,p_r,x,p_l}(t_0) \cdot S^4_{y,p_r,d_0}(t_0)}^{\text{Input: routing scheme}} \Bigg)$$
$$- \Bigg(\overbrace{\rho^4_{x,p_l}(t_0)}^{\text{Respiration rate}} + \overbrace{\kappa^{4,5}_{x,p_l}(t_0)}^{\text{Output: } c_5 \text{ pool}} + \overbrace{\kappa^{4,6}_{x,p_l}(t_0)}^{\text{Output: } c_6 \text{ pool}} + \overbrace{\sum_{y \in Q(x)} P^+_y(t_0) \cdot \tau_x}^{\text{Output: routing scheme + extra respiration}}$$
$$+ \overbrace{\lambda_{x,p_l,d_0}(t_0)}^{\text{Output: erosion t} \rightarrow \text{b}} + \overbrace{\sum_{y \in Q(x)} \sum\nolimits_{r=1}^{v} \zeta_{y,p_r,x,p_l}(t_0)}^{\text{Output: vertical flow t} \rightarrow \text{b}} \Bigg) \cdot \overbrace{S^4_{x,p_l,d_0}(t_0)}^{c_4 \text{ stock}} \tag{2.13}$$

For the other pools c_5 and c_6, the equations are analogous.

Middle layers The routing scheme for valley bottoms also affects the middle layers with its vertical components. For having layers above and below, such layers have fluxes in both directions. For a PFT p_l of pool c_4, the equation for d_k, $0 < k < m$, is:

$$\frac{dS^4_{x,p_l,d_k}(t_0)}{dt} = \Bigg(\overbrace{\sum_{c_i \in C_l} I^{*4}_{x,p_l,d_k}(t_0)}^{\text{Input: litter pools}} + \overbrace{\kappa^{6,4}_{x,p_l}(t_0) \cdot S^6_{x,p_l,d_k}(t_0)}^{\text{Input: } c_6 \text{ pool}} + \overbrace{\kappa^{5,4}_{x,p_l}(t_0) \cdot S^5_{x,p_l,d_k}(t_0)}^{\text{Input: } c_5 \text{ pool}}$$
$$+ \left[\overbrace{\lambda_{x,p_l,d_{k-1}}(t_0)}^{\text{Input: erosion t} \rightarrow \text{b}} + \overbrace{\sum_{y \in Q(x)} \sum_{r=1}^{v} \zeta_{y,p_r,x,p_l}(t_0)}^{\text{Input: vertical flow t} \rightarrow \text{b}} \right] \cdot S^4_{x,p_l,d_{k-1}}(t_0)$$
$$+ \overbrace{\sum_{y \in Q(x)} P^+_y(t_0) \cdot \tau_x \cdot S^4_{x,p_l,d_{k+1}}(t_0)}^{\text{Input: vertical flow b} \rightarrow \text{t}} \Bigg)$$
$$- \Bigg(\overbrace{\rho^4_{x,p_l}(t_0)}^{\text{Respiration rate}} + \overbrace{\kappa^{4,5}_{x,p_l}(t_0)}^{\text{Output: } c_5 \text{ pool}} + \overbrace{\kappa^{4,6}_{x,p_l}(t_0)}^{\text{Output: } c_6 \text{ pool}} + \overbrace{\sum_{y \in Q(x)} \sum_{r=1}^{v} \zeta_{y,p_r,x,p_l}(t_0)}^{\text{Output: vertical flow t} \rightarrow \text{b}}$$
$$+ \overbrace{\lambda_{x,p_l,d_k}(t_0)}^{\text{Output: erosion t} \rightarrow \text{b}} + \overbrace{\sum_{y \in Q(x)} P^+_y(t_0) \cdot \tau_x}^{\text{Output: vertical flow b} \rightarrow \text{t}} \Bigg) \cdot \overbrace{S^4_{x,p_l,d_k}(t_0)}^{c_4 \text{ stock}} \tag{2.14}$$

Bottom soil layer: d_m Finally, for the bottom soil layer, the equation is identical as Equation 2.14, the exception being the inexistence of b $\rightarrow$ t input or t $\rightarrow$ b output rates, since there are no bottom layers. In this case, for a given PFT p_l of pool c_4, we have:

$$\frac{dS^4_{x,p_l,d_m}(t_0)}{dt} = \Bigg(\overbrace{\sum_{c_i \in C_l} I^{*4}_{x,p_l,d_m}(t_0)}^{\text{Input: litter pools}} + \overbrace{\kappa^{6,4}_{x,p_l}(t_0) \cdot S^6_{x,p_l,d_m}(t_0)}^{\text{Input: } c_6 \text{ pool}} + \overbrace{\kappa^{5,4}_{x,p_l}(t_0) \cdot S^5_{x,p_l,d_m}(t_0)}^{\text{Input: } c_5 \text{ pool}}$$
$$+ \left[\overbrace{\lambda_{x,p_l,d_{m-1}}(t_0)}^{\text{Input: erosion t} \rightarrow \text{b}} + \overbrace{\sum_{y \in Q(x)} \sum_{r=1}^{v} \zeta_{y,p_r,x,p_l}(t_0)}^{\text{Input: vertical flow t} \rightarrow \text{b}} \right] \cdot S^4_{x,p_l,d_{m-1}}(t_0) \Bigg)$$
$$- \Bigg(\overbrace{\rho^4_{x,p_l}(t_0)}^{\text{Respiration rate}} + \overbrace{\kappa^{4,5}_{x,p_l}(t_0)}^{\text{Output: } c_5 \text{ pool}} + \overbrace{\kappa^{4,6}_{x,p_l}(t_0)}^{\text{Output: } c_6 \text{ pool}} + \overbrace{\sum_{y \in Q(x)} P^+_y(t_0) \cdot \tau_x}^{\text{Output: vertical flow b} \rightarrow \text{t}} \Bigg) \cdot \overbrace{S^4_{x,p_l,d_m}(t_0)}^{c_4 \text{ stock}} \tag{2.15}$$

2.2.2 . Study area

In this work, the study area comprises the European Union member states (EU27), plus Switzerland, the United Kingdom, and the Balkan states (i.e., Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia and Serbia). The EU27 is a political and economic block of 27 countries, covering 410 million hectares – larger than the seventh-largest country in the world (India) – and 447 million inhabitants. Switzerland, the United Kingdom, and the Balkan States were included for being spatially adjacent territories. The food and farming sector of EU27 used 156.7 million hectares of land (i.e., 38.2% of the total area) for agricultural production in 2016 and currently provides nearly 40 million jobs (i.e., 9.75% of the total population) (Statistical Office of the European Union, 2020; European Union, 2021).

EU27 has been promoting changes to shift its agriculture towards more sustainable practices. In 2019, for example, the European Commission proposed the European Green Deal, a growth strategy for the continent that proposes environmental targets, including climate neutrality by 2050. Some of the targets include increasing the share of organic farming from 8.5% of the total agricultural land to 30% by 2030 and increasing tree cover by planting 3 billion additional trees also by 2030 (European Commission, 2021). Such actions come as an anticipated response to projections of future environmental conditions. For example, the projected patterns of rainfall erosivity for the future indicate an increase in 81% of the European territory by 2050 (Panagos et al., 2017), which will consequently affect soils, a very rele-

vant natural resource for the achievement of the European Green Deal's goals (Montanarella and Panagos, 2021).

Input data: LSM emulator and erosion rates

The first step to running CE-DYNAM is to build a standalone version of the soil carbon dynamics of an existing LSM, i.e., an emulator. Such procedure is done by carefully analyzing and modifying the source code of the original LSM to allow the export of all necessary variables for reproducing calculations externally. In this work, we ran ORCHIDEE, a process-based model that simulates vegetation, energy, water, and carbon fluxes (Krinner et al., 2005), with the following settings: i) version: ORCHIDEE 2.2; ii) time step and range: daily, from 01/01/1860 to 31/12/2018; ii) climate: monthly forcings at a 0.125° spatial resolution from the VERIFY project (see `https://verify.lsce.ipsl.fr/index.php/presentation`); iii) land cover: annual forcing – derived from the ESA CCI Land Cover dataset (European Spatial Agency, 2021; LSCE, 2021).

For the calculation of erosion rates, we applied the well-known Revised Universal Soil Loss Equation (RUSLE) model, using the values recently developed by the European Commission specifically for our study area (Panagos et al., 2015e). For a given year (y) and month (m), monthly erosion rate (E), in $t/(ha \cdot year \cdot PFT)$ was calculated as:

$$E(y,m) = R(y,m) \cdot K \cdot C(y) \cdot LS \cdot P \tag{2.16}$$

with $R(y,m)$ being the rainfall erosivity factor[1] in $MJ \cdot mm \cdot ha^{-1} \cdot h^{-1} \cdot yr^{-1}$, K being the soil erodibility factor[2] in $t \cdot ha \cdot h \cdot ha^{-1} \cdot MJ^{-1} \cdot mm^{-1}$, $C(y)$ being the dimensionless land cover and management factor[3], LS being the dimensionless slope length and steepness factor[4], and P being the dimensionless support practices factor[5]. When collapsing the PFT dimension for

[1] It represents the "kinetic energy of raindrop's impact and the rate of associated runoff" (Panagos et al., 2015a; Wischmeier et al., 1978)

[2] It "represents an integrated annual value of the soil profile reaction to the process of soil detachment and transport by raindrops and surface flow" (Panagos et al., 2014; Renard, 1997)

[3] It "accounts for how land cover, crops, and crop management cause soil loss to vary from those losses occurring in bare fallow areas" (Panagos et al., 2015c; Kinnell, 2010)

[4] It describes "the effect of topography on soil erosion" (Panagos et al., 2015b)

[5] It "accounts for control practices that reduce the erosion potential of runoff by their influence on drainage patterns, runoff concentration, runoff velocity and hydraulic forces exerted by the runoff on the soil surface" (Panagos et al., 2015d; Renard et al., 1991)

the calculation of annual or monthly averages, $E(y, m)$ was multiplied by the corresponding land cover fraction p_i for PFT i (Table 2.1). All the data was aligned and processed on a 0.125° grid.

As seen in Equation 2.16, factors K, LS and P were assumed constant for the whole period 1860 – 2018, while $R(y, m)$ varies per month and $C(y)$ varies annually. The source of K is the extrapolated version of Panagos et al. (2014) including stoniness, LS comes from the completely harmonized version of Panagos et al. (2015b) for the whole study area, and P comes from the database provided by Panagos et al. (2015d). The first and the second factors cover the whole study area originally, but the third does not, so an additional assumption was added: in places where P was not available (i.e., Switzerland and the Balkan states), it was assumed to be equal to 1. According to the authors mentioned above, K and P were derived from field survey data carried out in 2009 and 2012, respectively.

For C, we used the spatial dataset of Panagos et al. (2015c), but an additional procedure was made to minimize the differences arising from the mismatch in the land cover classes definition and spatial resolution. Such a procedure consisted of fitting a linear regression model to the upscaled version of the original C factor using the target land cover classes as explanatory variables (i.e., $C = \sum_i \beta_i \overline{p}_i + \epsilon, \quad \epsilon \sim \mathrm{N}(0, \sigma^2)$). An intercept term was intentionally not added to the linear regression, and $\overline{p_i}$ is the average land cover from the period 2010-2018, approximately the period of data collection of Panagos et al. (2015c).

The rainfall erosivity also demanded an extra processing step. The main source for calculations was the monthly erosivity derived and provided by Ballabio et al. (2017). To extrapolate for the past, we assumed a constant erosivity density for the whole simulation period, 1860-2018. That was made by calculating:

$$R(y, m) = r(y, m) \cdot \frac{R^*(m)}{\overline{r}(m)}$$

with $R^*(m)$ being the original monthly erosivity data set upscaled to a 0.125° spatial resolution, $\overline{r}(m)$ being the average monthly precipitation of the period 2010-2018 (roughly the same data collection period of Ballabio et al. (2017)), and $r(y, m)$ being the monthly precipitation for the month m of year y.

Calibration and validation

Calibration Calibrating CE-DYNAM means ensuring that the values predicted by the routing scheme introduced are realistic and consistent with field observations. To do so, we made an exact copy of the model described in Equation 2.2.1 and Equation 2.2.1, but replacing the carbon quantities with sediment quantities. Then, we used as field data the information of total suspended solids and river discharge from the GEMStat database (United Nations Environment Programme, 2018b) for the whole Europe. We adopted a squared error cost function between the model predictions and the observations. Because the calculation of analytical derivatives of the cost function with respect to the parameters is hard in our case, minimization was performed using the NEWUOA solver (Powell, 2006, 2008) with early stopping to prevent overfitting.

Like most optimization methods, NEWUOA requires several evaluations of the cost function, which is computationally expensive in our case. For this reason, we calibrated the model using annual averages instead of monthly inputs to accelerate calculations. We also preprocessed our observations by first aggregating annually the total of 10,552 instantaneous observations available, which resulted in 391 annual median values for 40 rainfall stations distributed across Europe from 1979 to 2003. Then, we calculated the 5-year moving averages of the median annual values to simultaneously smooth extreme values from floods that are not modeled in CE-DYNAM and remove stations with a few observations. The final dataset contained 241 observations in 30 stations, whose contributing areas covered nearly one-fourth (23.34%) of the study area. For each set of parameters, we calculated the predicted sediment stock in the river fraction of the cell (from variable l_x, see Table 2.1), and the objective function adopted was the squared error between this quantity and the product between the total suspended sediments observed and the water volume in a day (derived from the instantaneous discharge). As in similar works such as Borrelli et al. (2018b), the model was assessed using the Nash-Sucliffe model efficiency (ME) coefficient (Nash and Sutcliffe, 1970), and the coefficient of determination (R^2) as defined by Everitt and Skrondal (2010).

Validation The hillslope erosion rates were externally validated by comparing our estimates with some field observations and modeled values in the literature. We used the compilation of observations by Cerdan et al. (2010) in two ways: i) we aggregated our and their land cover classes into four common categories (i.e., croplands, grasslands, forest, and bare soil), and compared the distribution of our calculations with their reported point estimates; ii) we

compared our country averages with their extrapolated calculations for the whole Europe. We also compared our erosion values to the compilation of local-scale field observations from different sources reported by Panagos et al. (2020) and modeled country averages of (Panagos et al., 2015e) to evaluate the model behavior at a local and regional scale, respectively.

The calibration of the lateral movements on the model was validated both internally and externally. The model's internal consistency was checked by comparing the physical quantities obtained empirically by the calibration procedure with the results previously obtained in the literature. We checked the model's ability to predict the sediment concentrations in places unobserved during the calibration process. For this purpose, the 30 rainfall stations of the final dataset were divided between 24 stations for calibration - i.e., 80% observed by the model - and six stations for validation - i.e., 20% unobserved by the model.

2.2.3 . Simulations

In order to evaluate the behavior of CE-DYNAM under different scenarios, two simulations were made after model calibration. In Simulation #1 (S1), all ETD-related processes are considered. In Simulation #2 (S2), no ETD-related processes are added to the original LSM fluxes. In this case, the original model is only affected by the vertical discretization of fluxes and the division of soil carbon pools into hillslopes and valley bottoms soil carbon pools. With such assumptions, the summation of all the results for soil layers of S2 recovers the original LSM results. In both cases, simulations were run from 01/01/1860 to 31/12/2018.

2.3 . Results and discussion

2.3.1 . LSM emulator and erosion rates

The main result for the LSM emulator is presented in Figure 2.4: a comparison of the true values of ORCHIDEE against the predicted ones from the emulator with ETD not enabled. The simulation was made in one random grid cell representative of the model behaviors over the entire studied region. A slight mismatch between original and predicted values exists at early timesteps, but as expected, values tend to a nearly identical curve after a few timesteps, indicating the adequacy of the emulator to replace the full LSM for erosion calculations with CE-DYNAM. In general, the adoption of an emulator has advantages and disadvantages for

CE-DYNAM compared to its implementation directly on an LSM. On the one hand, one can list its simplicity, agility, and flexibility as an advantage to be easily modified for the inclusion of new dynamics, such as the ETD fluxes in the present research, or for other LSMs. It could be noted at this point that practically all existing soil carbon model implementations can be represented in a linear form, and therefore, the matrix approach could be applicable and CE-DYNAM, coupled (Huang et al., 2018). In fact, Sierra and Müller (2015) and Metzler et al. (2020) demonstrate how the approach could be used even for more complex nonlinear models. On the other hand, the use of a standalone version of the LSM allows the processes to be represented only in a simplified way. In the present work, for example, the respiration and litter input rates simulated by the LSM are unchanged in CE-DYNAM, whereas the literature suggests that, in fact, they should also be affected by ETD fluxes (Olson et al., 2016). Such a limitation also exists for other important interactions affecting the fate of transported carbon that cannot be properly incorporated into the emulator, such as variation in soil moisture and temperature, as well as in organic matter quality and soil fractions (Lal, 2003).

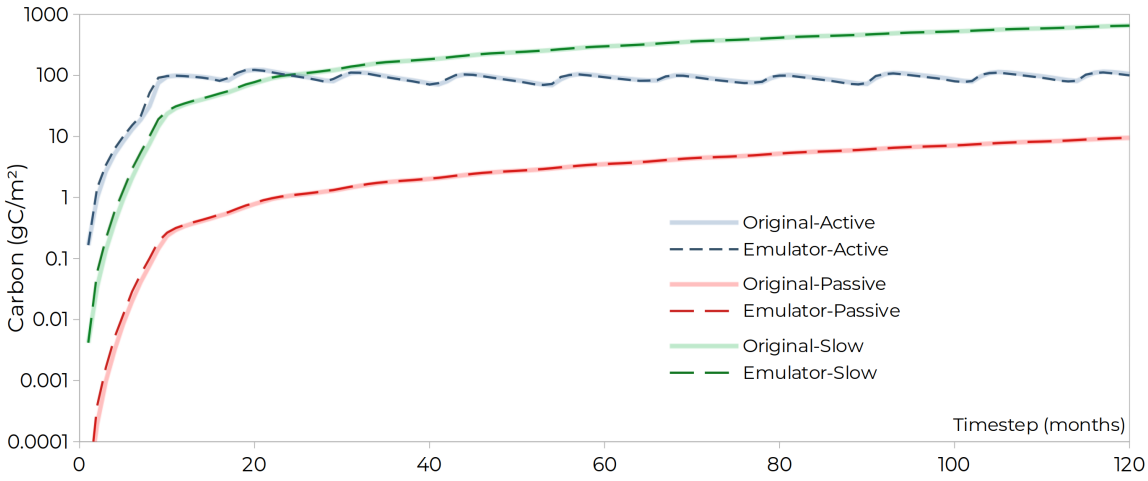


Figure 2.4: Comparison between the results of the original LSM (i.e., ORCHIDEE, continuous lines) and the results of the emulator constructed for the present work (dashed lines), for 10 years of pools initialization. ORCHIDEE does not have a hillslope-valley bottoms differentiation of pools, while the line for the emulator corresponds to their sum. Ideally, the emulator would be a perfect standalone version of the LSM.

For the erosion rates, Figure 2.5 and Figure 2.6 present the historical spatio-temporal variability reconstructed. On the top subfigure of Figure 2.5, the absolute erosion rates in 1860 are shown on the left, while the maps for 1910, 1960, and 2010 represent the difference

with respect to 1860. It can be seen that the annual variations do not follow a linear pattern through time, thus affecting erosion calculations unequally. A decrease in erosion rates from 1860 to 2010 can be noted in Central Europe, and a strong pattern on countries' borders follows from the assumptions of the reconstructed land cover database used (LSCE, 2021). Although partially, this result is related to those reported by Bork and Lang (2003) and Dotterweich (2008), who in historical reconstructions in Germany and Central Europe found peaks in erosion rates in the second half of the 18th century, a period for which there is documentary evidence of extreme rainfall events, and in the early 19th century. In Figure 2.5, the 1860 erosion map also shows points in four different locations (i.e., P1, P2, P3, and P4, plus the whole study area (WSA)). Within a year, the monthly variations in erosion rates are due solely to changes in the rainfall and the erosivity factor, and as seen on the bottom graph of Figure 2.5, the pattern of such changes also vary non-linearly in space and may differ from the average pattern of the study area. Additionally, Figure 2.6 shows the annual average erosion rates, calculated as 2.96 t/ha for 2018. Such a value is higher than the 2.46 t/ha reported by Panagos et al. (2015e), which can be justified by the different spatial resolution, land cover database, and study area since, in our case, we include Switzerland and the Balkan states, which have erosion rates that are relatively higher compared to their neighbors (Figure 2.5, top left). Figure 2.6 also shows the average effect of the two time-varying factors adopted for RUSLE calculation, i.e., the R- and the C-factors of Equation 2.16. Two distinct patterns of variation can be seen through time, with the R-factor having a higher annual variability and the C-factor being less abrupt, except for the evident breaks during the two World Wars. The R-factor shows a cyclic pattern from 1860 to 1901 due to the recycling of (i.e., repetition of climate) forcings adopted by ORCHIDEE for this period. The calculations also indicate an overall increase in erosion rates from 1860-to 1960 due mainly to land cover changes and a peak in rainfall erosivity, followed by an overall decrease from 1960-to 2018. Despite such a pattern in the nearer past, the literature indicates a tipping point in the present. Future projections from Panagos et al. (2021) indicate that water erosion in Europe is expected to increase between 13 and 22.5% by 2050, and Borrelli et al. (2020) estimate an increase in 2015 water erosion rates of 33 to 66% by 2070 worldwide. Under these scenarios, future values could be even higher than the past values calculated and shown in Figure 2.6.

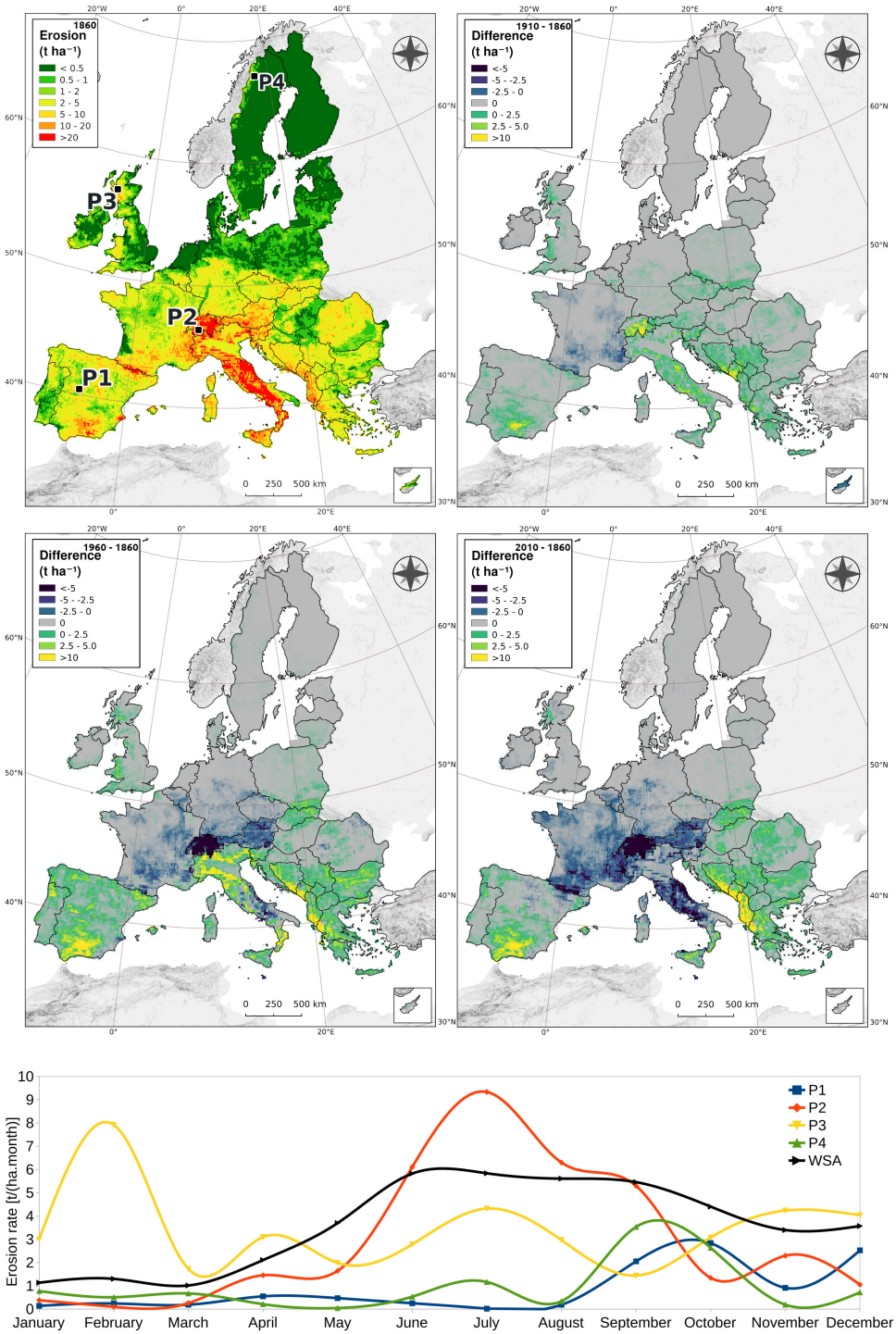


Figure 2.5: Erosion rates calculated in this work. On the top subfigure, the top-left map shows the erosion rates for 1860 and four points (P1, P2, P3, and P4), while the top-right, bottom-left, and bottom-right maps show the anomalies for 1910, 1960, and 2010, respectively. The bottom subfigure shows the changes in erosion rates due to variations in the monthly rainfall and erosivity for P1, P2, P3, P4, and the whole study area (WSA) within the same year, 1860.

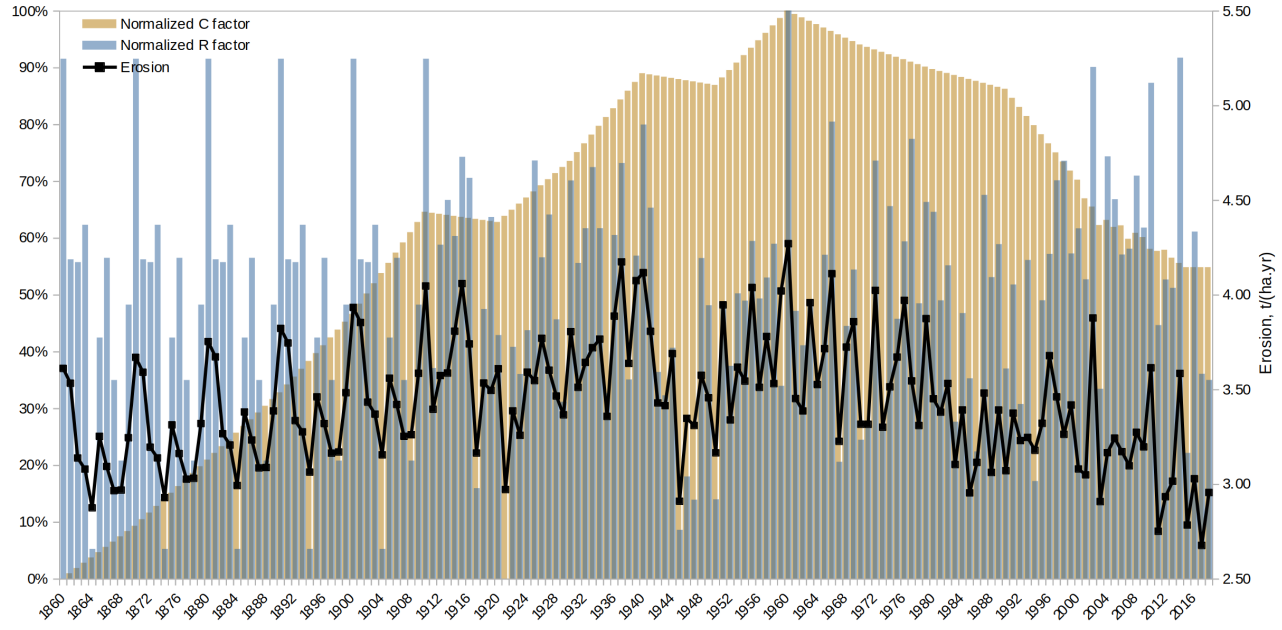


Figure 2.6: The average impact of the reconstructed R and C-factors on the erosion rates for the period 1860-2018.

Also for the erosion rates, the annual country averages were compared against values reported in the literature. The results of Figure 2.7 (left) show a positive agreement between all databases considered, as highlighted by the positive slope of the robust linear models fitted to the data. The steepness of curves suggests that the model's ability to reproduce the continent-scale patterns is higher than local-scale predictions, which can be interpreted as a consequence of the model's relatively coarse resolution to represent local-scale hydrology. On the right part of Figure 2.7, the comparison per land cover class shows a close match for croplands and bare soil, the highest rates in our model. On the other hand, our model tends to underestimate erosion in forests and grasslands compared to the external sources, with our values lying on the lower tail of the distribution of Panagos et al. (2015e) and Cerdan et al. (2010).

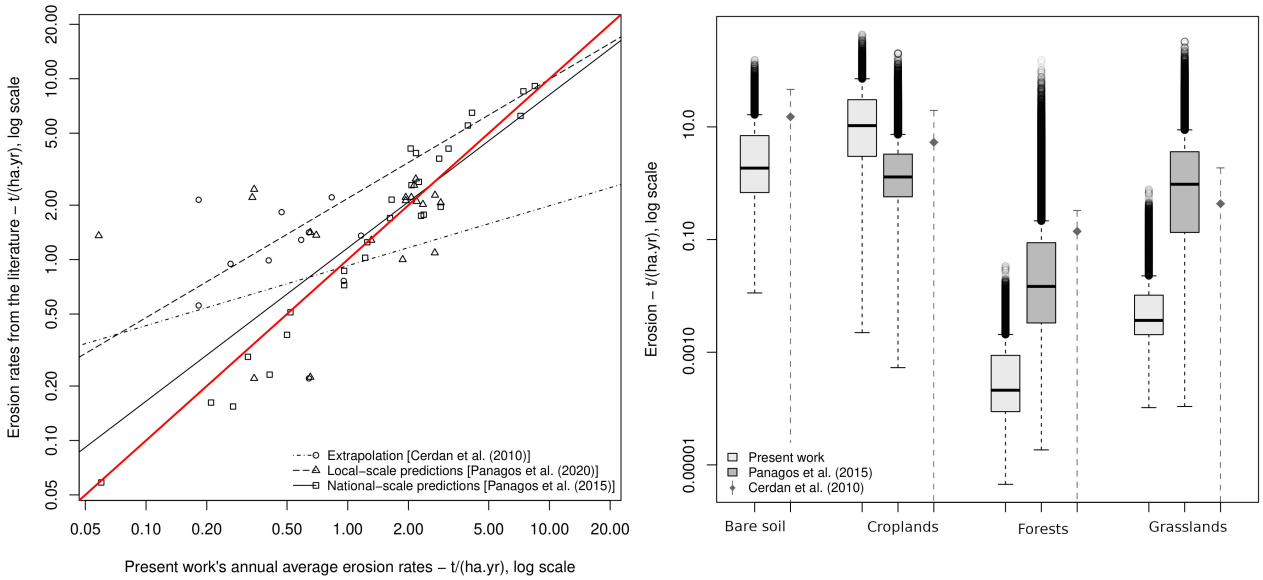


Figure 2.7: Comparison between the average erosion rates calculated by external sources versus the values calculated in our work, along with an identity line (left); and the comparison of the distribution of erosion rates per land cover (right). In the right plot, the values for our work are the average for the period 1970-2018, and the values for Cerdan et al. (2010) are the reported mean $\pm$ standard deviation.

2.3.2 . Model calibration

The NEWUOA algorithm performed several hundred function evaluations until stop. Using averaged annual instead of monthly forcings and performing the calculation only for the catchments areas of stations, each evaluation took around 6 minutes of processing time. The vertical discretization parameter obtained was $\gamma_1 = 0.1$, approximately the left pattern from Figure 2.2. Calibration also indicated that the input of sediments from hillslopes into valley bottoms varies from 0.4 to 11.8% in croplands, from 4.9 to 10.9% in forests, and from 0.3 to 3.8% in grasslands (Figure 2.8, left). These values can be interpreted in several ways. First, the absolute magnitude of the values is relatively small, following what was suggested by Hoffmann et al. (2013a), with the maximum values being comparable to the 15% of on-site erosion reaching riverine systems presented by Borrelli et al. (2018b) for the same study area. However, the direct comparison of these values should be read with caution because of the large methodological differences between works. For example, the authors defined the rivers explicitly and used a higher spatial resolution for a single moment in time, characteristics that contrast with those of CE-DYNAM (Naipal et al., 2020). Furthermore, despite

the similar interpretation, the quantities compared may themselves differ between the models adopted (Rompaey et al., 2001). Second, regarding the shape of the curves, there is an increasing relation in parameter g_f as a function of slope for forests and croplands but a decreasing relation for grasslands. The increase in two of the three land-use classes can be readily explained by the important effect of gravity on sediment transport (Bridge, 2003; Huggett, 2017), while its generally low range of values can partially explain the unexpected decrease in grasslands compared to that of forests and croplands. Third, with respect to the ordering of the curves, two patterns are observed. In flat areas, with a slope less than 1.5°, the pattern is forests having higher transport than grasslands, followed by croplands. In steeper areas, with slopes above 1.5°, there is a rapid change in the ordering, leading to a situation where croplands generally have higher sediment transport than forests and grasslands. The low influence of this region on sediment production can explain the non-intuitive relationship between the classes. For example, areas with slopes less than 1.5° were responsible for only 5.43% of Europe's total erosion in 2018, with the remaining 94.57% occurring in the steeper areas. Therefore, it is reasonable to expect that the steeper areas will be better represented in the model. Thus, the latter pattern of Figure 2.8 can be explained by the lower cohesive properties of the less vegetated covers relative to the more vegetated ones, consequently offering less resistance to water and sediment flow (Osterkamp et al., 2011; Hoffmann et al., 2013a; Huggett, 2017). Figure 2.8, right, also shows the sediment residence time, which was estimated to vary from 0.5 years (180 days) to 24.5 years, indicating that sediment retention increases with watershed size, in agreement with that described by Hoffmann et al. (2013b).

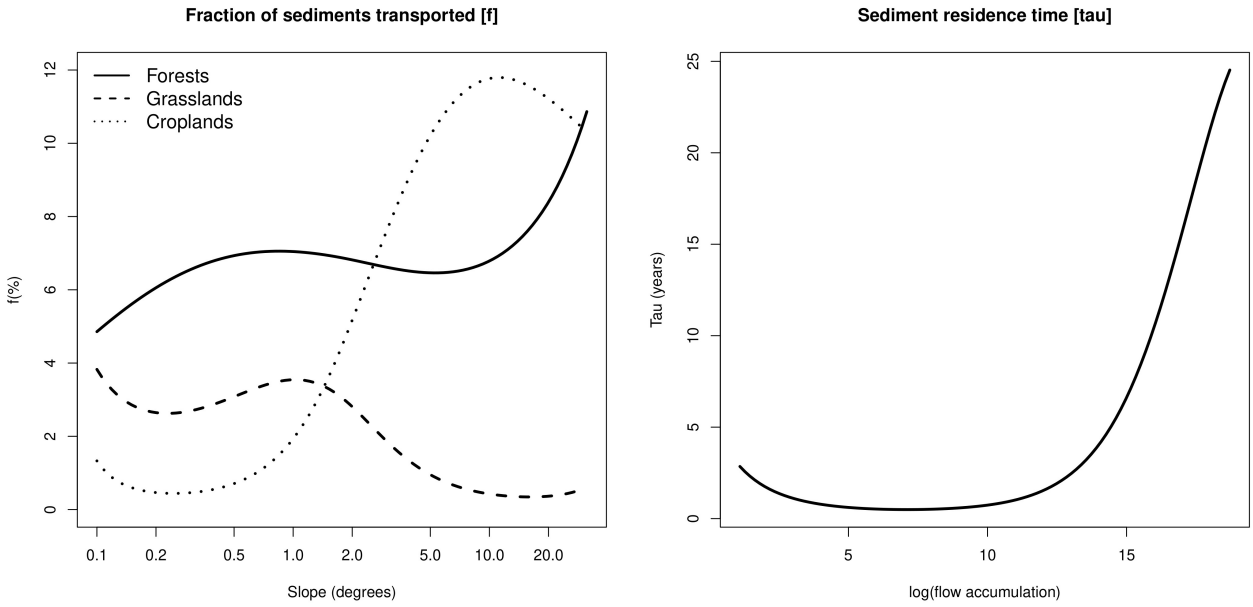


Figure 2.8: Results from the calibration procedure: the fraction of sediments that go from hillslopes to valley bottoms (left), and sediment residence time (right).

The validation for the best set of parameters is summarized in Figure 2.9, which shows a plot of actual against predicted values (using annual forcings) for the observed and the unobserved locations. In all cases, the prediction value used for this comparison are those from the same year as the observation. The red diagonal is a 1:1 line, and each group of dots connected by a line represents a different station. The best set of parameters found yielded ME $= 0.603$ and $R^2 = 0.666$ for the observed stations, and ME $= 0.152$ and $R^2 = 0.438$ for the unobserved stations. Overall, for the full dataset, the model has a ME $= 0.578$ and $R^2 = 0.640$. The values obtained are relatively high when compared to similar studies. Works such as that of Feng et al. (2010) and Rompaey et al. (2005) in China and Italy, respectively, reported negative ME values, which indicate that sometimes distributed models are unable to represent sediment dynamics, especially when there is high heterogeneity in the data (Rompaey et al., 2005). In Quijano et al. (2016), where the authors studied four adjacent hydrological units at a local scale in Spain, distributed models could represent well the dynamics involved. The overall value obtained for the study region was ME $= 0.11$, while the value calculated individually per hydrological unit ranged from ME $= -0.11$ to ME $= 0.49$. In what is probably the most similar to our work in terms of the study area, Borrelli et al. (2018b) initially con-

sidered a total of 24 semi-natural and agricultural basins in Europe, for which they obtained ME = 0.38. The result motivated the authors to further remove basins as a fine-tuning of the model calibration used, which led to a ME $= 0.89$ for ten basins. The results from these other studies help us to compare the performance of the model presented in the current work. It can be noted that the present work uses more observations and calibrates the model with time-varying data (i.e., not long-term averages), which requires a more complex model architecture and highlight the robustness of the calibration performed. It is also an important remark that the comparison with other works was only possible after the methodological improvements in the new version of CE-DYNAM compared to that of Naipal et al. (2020): i) the possibility of calibrating the lateral fluxes using sediment data collected in the field and relatively abundant in the literature (see United Nations Environment Programme, 2018b); ii) the possibility of performing validation with field data, using the model as a basis for prediction for locations unobserved during its fit.

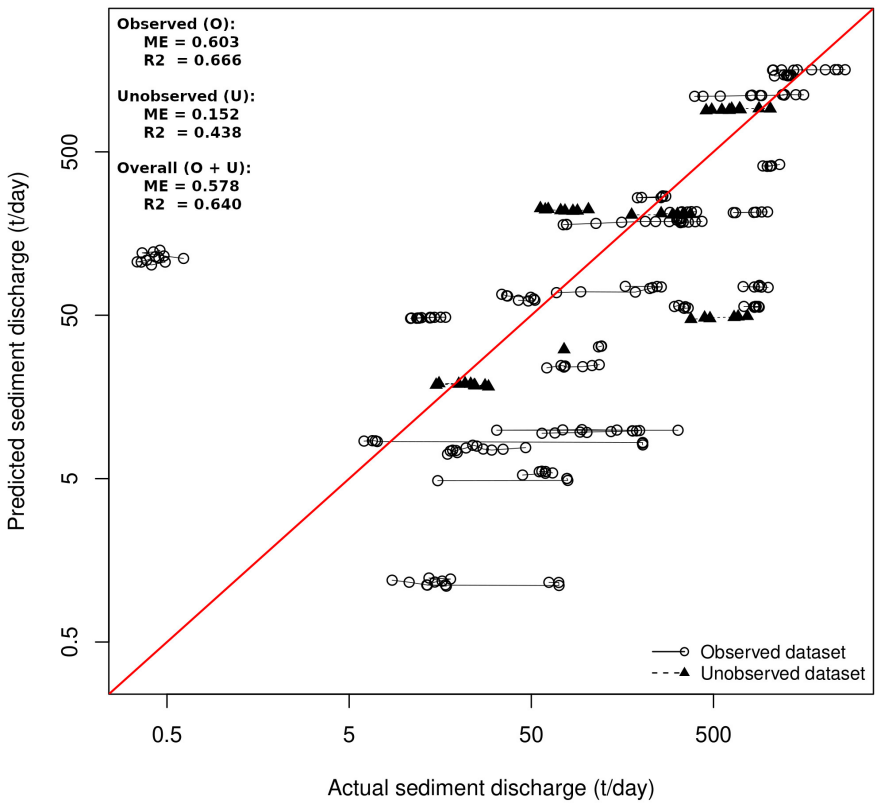


Figure 2.9: Plot of the predicted annual averages against the observed annual sediment discharged values in log scale. Each dot or triangle corresponds to one of the 241 observations, and the connected icons correspond to one of the 30 different stations of the database. The red line is an identity line. The unobserved dataset contains the observations not included in the model calibration procedure.

We also used the same best set of parameters with monthly forcings to quantify how distant the simplified calibration with annual forcings is from the optimal condition. Using monthly forcings for a full calibration remains precluded, since a single function evaluation took almost one day to complete. In that experiment, values dropped to ME $= 0.464$ and $R^2 = 0.616$ for the entire dataset, indicating a relatively small change in predictions compared to the simplification using annual forcings.

2.3.3 . Model behavior under the matrix approach

In both scenarios, the two A matrices (i.e., one for hillslopes and the other for valley bottoms calculation) are square with a size equal to the product between the number of cells, the number of PFTs, the number of soil layers and the number of soil carbon pools, which for the setting used in the present work equals $1.85 \cdot 10^7$ rows and columns. In S1, where all fluxes are considered, the average number of non-zero elements on the A matrices of hillslopes and valley bottoms were $7.4 \cdot 10^6$ and $1.7 \cdot 10^7$, respectively, corresponding to densities of 21.6 ppb (i.e., 10^{-9}) and 49.7 ppb. In S2, where fewer fluxes are simulated, the A matrices for hillslopes and valley bottoms contained $5.9 \cdot 10^6$ and $7.4 \cdot 10^6$, yielding a density of 17.2 ppb and 21.6 ppb, respectively. Simulations were run on a High-Performance Computer having 19 Intel(R) Xeon(R) CPU E5-2640 v4 2.40 GHz processors. The number of cores used for calculations varied, as specified next. In both scenarios, the generation of A and B took around 3 to 6 minutes for each simulation month at the continent scale. After the generation of all matrices, equilibrium calculation took around 10 minutes on a single core. As a comparison, the calculation for the non-alpine region of the Rhine basin of Naipal et al. (2020) used to take two days in seven cores of the same machine. The drastic reduction is a consequence of the matrix approach adopted. While in their work all fluxes had to be recalculated for each month of equilibrium calculation, in our case we just had to precalculate A and B once.

The matrix approach brought several benefits for speeding up the model. First, the analytical representation of the model allowed the derivation of first-order approximations for the monthly averages, which are faster to calculate than summing the daily simulations and dividing by the number of days. Second, despite the different settings of S1 and S2, implementation was straightforward thanks to the natural interpretation of each element of A as a flux of carbon from one to another uniquely identified combination carbon pool. Third, the calculation of A and B are independent for each month, which drastically increases the number of possible concurrent execution threads in comparison to the original model of Naipal et al. (2020). These results are similar to those found in the literature. For example, Xia et al. (2012) and Huang et al. (2018) reported reduced processing time and computational cost of the calculations. In our case, it is not possible to compare the processing times as done by Xia et al. (2012) because the previous version of the CE-DYNAM from Naipal et al. (2020) does not support the calculation at a continental scale in a feasible time. In this sense, the very possibility of applying the model at this scale, now allowed because of the matrix approach, is an indicator of such improvements. However, the matrix approach may require changes for applications at even larger scales, such as the global scale. As shown in Equation 2.2, the approach simplifies the simulation by representing the variation of fluxes with additions and

matrix-vector multiplications. Such operations are typically efficient in the format known as Compressed Sparse Row (CSR) (Bai, 2000; Greathouse and Daga, 2014), which requires the storage of three vectors for its construction. In global scale problems, it is possible that the amount of memory required for such storage is excessively large, so alternative representations should be explored. A possible solution may come from an analogous problem in the statistical regression literature, where authors seek low-rank representations of models in order to reduce the dimensionality of the problem while still largely preserving the characteristics of the original model (see Wang and Ranalli (2006) and Wood (2006) for two examples). However, while procedures that are central for dimensionality reduction problems such as singular value decomposition are well established for dense matrices (Strang, 2016) or even sparse matrices of reasonable size, the problem can be complicated when the dimensions are huge. Therefore, further research and work to search for applicable methods are needed.

The calculation for a depositional site indicates that the additional incoming fluxes from its upstream area due to ETD processes tend to increase the equilibrium carbon stock at the site from 6,800 to 7,150 g/m^2, equivalent to a 5.1% increase (Figure 2.10, top). The relative increase compared to the equilibrium stock showed in Figure 2.10 bottom, illustrates the ability of the model to emulate non-linearities on the impact of ETD fluxes on the carbon cycle at a depositional area. According to van Oost et al. (2005), Li et al. (2007), and Wang et al. (2015), some expected changes in the dynamics include an increase in the C burial, resulting in an increase of the soil organic carbon, as well as enhanced respiration of the carbon buried with time (Naipal et al., 2020). In fact, in the early period of the curve, from 1860 to approximately 1940, the curve of the S1 simulation (i.e., with ETD fluxes) is above that of the S2 simulation, indicating that the immediate impact of adding new fluxes to an area is to increase the rates of carbon burial. However, in the final period of the curve, from 1941 to 2018, the variation curve of S1 moves below the S2 curve (i.e., without ETD fluxes), indicating a decrease in the lateral input and the rate of carbon burial, as well as a higher respiration by microorganisms due to the carbon previously buried.

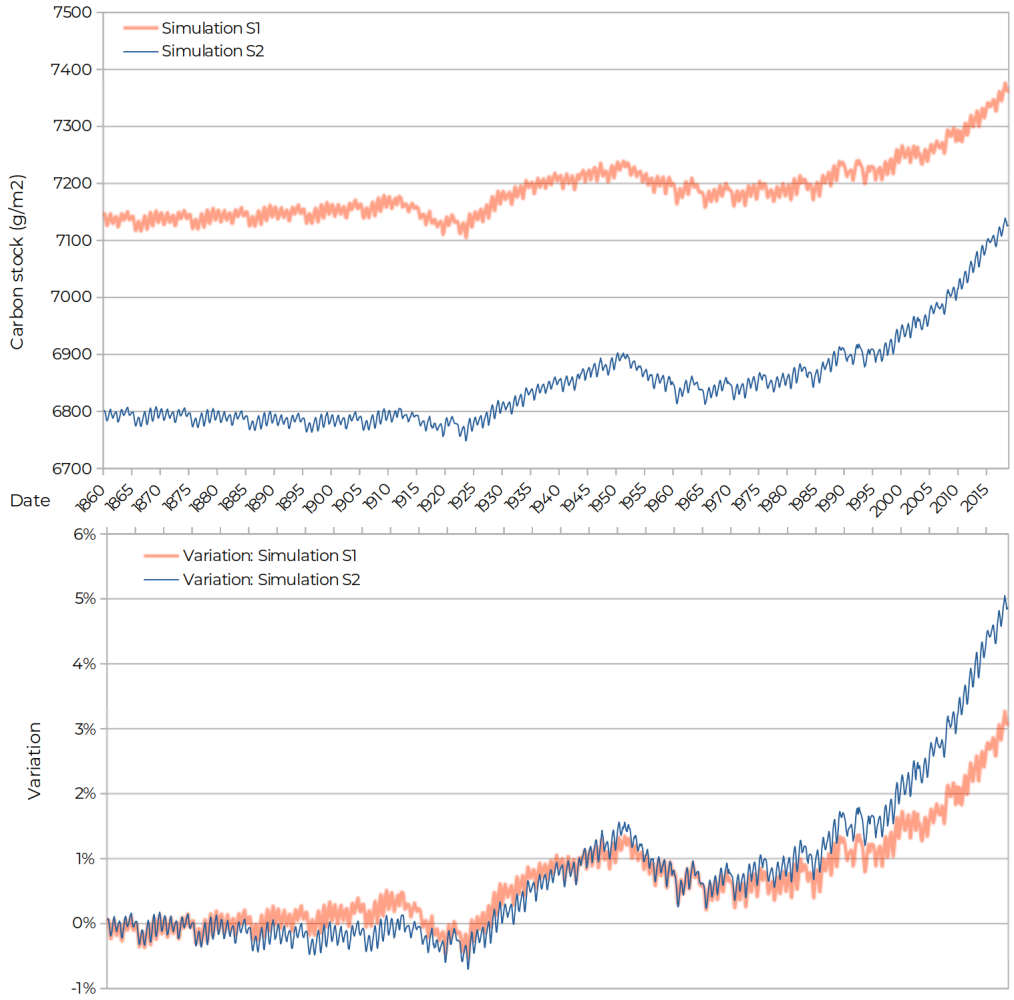


Figure 2.10: Results for the depositional site: absolute values of carbon stock (g/m^2) (top), relative difference compared to the stock at the first simulation month (bottom). Scenario S1 contains all ETD fluxes, while scenario S2 represents the dynamics of the original LSM.

Furthermore, to check the mass balance closure, all the fluxes from this depositional area were calculated from 1860 to 2018. On the cell's hillslope fraction, litter input added 25,404.47 gC/m^2 and land cover added 128.71 gC/m^2, of which 25,286.00 gC/m^2 were respired, and 0.39 gC/m^2 were sent to the valley bottoms. On the valley bottom fraction, the inputs were: 0.39 gC/m^2 coming from the hillslopes, 722.16 gC/m^2 from upstream lands, and 6,039.23 gC/m^2 from litter. Of this, 6,764.07 gC/m^2 was respired, and 23.91 gC/m^2 was lost due to land cover change. These values indicate that local erosion at this depositional area is not relevant to the carbon cycle, in contrast to the carbon input from upstream lands, which corresponded

to 10.68% of all the valley bottom fraction inputs.

2.4 . Limitations

Despite the advances presented in this work, there are still limitations that need to be addressed with future modifications. Concerning its structure, incorporating sediment data in CE-DYNAM is only possible by assuming that sediment and carbon follow the same dynamics. While this is convenient for calibration, it might not be realistic, so further work is necessary to improve this important assumption. Besides, the physical representation of the fluxes could be improved. For example, no transformation of C pools during the transport process is represented, such as the breakdown of aggregates. In practice, they could increase the turnover rates of soil organic carbon compared to the simulations we presented.

Regarding the calibration presented here, other limitations concern the spatial resolution and historical reconstructions. First, finding the optimal resolution for CE-DYNAM will always be a problem, since it is halfway between the fine-scale hydrological processes it represents and the coarse resolution of the current climate models. Second, the historical reconstructions presented are highly sensitive to the assumptions adopted and presented. Even though these assumptions are properly evaluated during the calibration and validation process, better results will be possible the better the input maps for the model are.

Another important limitation refers to the structure of the calibration parameters. The structure presented in this work contains a certain tradeoff between the share of sediments that move from hillslopes to valley bottoms versus the sediment residence time, so the optimization may sometimes tend to yield physically unrealistic results. In the present work, we attempted to solve this issue by using multiple starting points to the optimization procedure. In a more robust implementation, a better solution should come from the development of the model itself, which should mathematically penalize solutions that do not have a meaningful physical interpretation. Finally, our model also naturally inherits the problems of some necessary assumptions. Even the fundamental assumption that calculations converge to the pullback attracting trajectory in 1860-1869 might not be correct and affect the results largely (Sanderman et al., 2017; Dimassi et al., 2018). Analogously, the RUSLE-based approach for erosion modeling is widely criticized, but remains the sole alternative for large-scale quantitative applications (Panagos et al., 2016).

2.5 . Conclusions

In this work, we addressed the challenge of scaling CE-DYNAM, an erosion-transport-deposition model, in a relatively high spatial resolution and long period at the European scale. First, we show how the lateral fluxes of CE-DYNAM can be represented in a matrix form, an alternative that allows the acceleration of the computations performed, making them feasible for large-scale applications. Our work, therefore, enabled the previously precluded possibility of applying CE-DYNAM to large spatial domains or high spatial resolutions. We also improved the model's physical representation of sediment movement to allow for proper calibration and validation procedures using observations of sediment discharge collected on the field. With these changes, the presented model can be readily adapted to other study regions, the main limiting factor being the availability of inputs from external sources. We also describe how the proposed technical solution might not work on an even larger scale (global scale, for example), so further work may be needed to improve the proposed approach, such as the search for computation of low-rank representations of the model matrices.

Second, a more practical contribution of our work was the calibration of the model for the whole of Europe from field-collected data. Our results show that the patterns obtained are internally consistent and coherent with those previously reported in the literature in similar work. We expect the patterns obtained in this work to serve as a reference for future models for this study region. Since the calibration of the lateral fluxes is done using sediment data, the results form the basis for simulations of the impact of erosion on the carbon cycle and the future incorporation of other nutrients, such as nitrogen and phosphorus, into CE-DYNAM. These works could advance our understanding of the role of ETD processes on nutrient cycles.

Third, we used the calibrated model to predict the movement of carbon at a depositional site, the type of site that tends to be highly affected by incoming lateral fluxes from upstream lands. This simulation evaluated the model's impact on soil carbon pools and showed how the effect of erosion on the carbon cycle could be nonlinear in time. In this sense, this result shows that time-static models can only partially disclose the correct effect of ETD on the carbon cycle.

2.6 . Learnings

The first version of CE-DYNAM had design limitations that did not allow its application for large domains or higher resolutions. In this chapter, we focused on loosening these restrictions by formalizing the model equations and representing the lateral fluxes in matrix form. With these changes, we found that:

- Representing CE-DYNAM fluxes in matrix form simultaneously decreased its required runtime and allowed it to run at a continental level.
- The comparison between the erosion rates in CE-DYNAM showed a positive correlation with other sources measured or calculated at different scales.
- The historical reconstructions indicated that erosion rates increased from 1860 to 1960, followed by a decrease until the current years. Land cover and rainfall were found to drive long-term changes and short-term variability, respectively.
- A calibration procedure was introduced to overcome the imposed parameters in CE-DYNAM (v1).

The technical advances made in this chapter consisted of a necessary but not sufficient condition to use CE-DYNAM as a practical tool. At this point, CE-DYNAM (v2) could be run at a continental scale, but the physical representation of fluxes in the model needed further improvements. In the next chapters, we focused on some of these improvements: generating spatially-explicit cover crops information (Chapter 3) and redesigning the model structure and the calibration procedure (Chapter 4).

3 - From regional to parcel scale: a high-resolution map of cover crops across Europe combining satellite data with statistical surveys

This chapter was written by Arthur Nicolaus Fendrich, with contributions from Francis Matthews, Elise Van Eynde, Marco Carozzi, Zheyuan Li, Raphael d'Andrimont, Emanuele Lugato, Philippe Martin, Philippe Ciais, and Panos Panagos.

The reformed Common Agricultural Policy of 2023-2027 aims to promote a more sustainable and fair agricultural system in the European Union. Among the proposed measures, the incentivised adoption of cover crops to cover the soil during winter provides numerous benefits such as improved soil structure and reduced nutrient leaching and erosion. Despite this recognised importance, the availability of spatial data on cover crops is scarce. The increasing availability of field parcel declarations in the European Union has not yet filled this data gap due to its insufficient information content, limited public availability and a lack of standardization at continental scale. At present, the best information available is regionally aggregated survey data, which although indicative, hinders the development of spatially accurate studies. In this work, a statistical model is proposed relating Sentinel-1 data to the existence of cover crops at the 100-m spatial resolution over the entirety of the European Union and United Kingdom and estimate its parameters using the spatially aggregated survey data. To validate the method in a spatially-explicit way, predictions were compared against farmers' registered declarations in France, where the adoption of cover crops is widespread. The results indicate a good agreement between predictions and parcel-level data. When interpreted as a binary classifier, the model yielded an Area Under the Curve (AUC) of 0.74 for the whole country. When the country was divided into five regions for the evaluation of regional biases, the AUC values were 0.77, 0.75, 0.74, 0.70, and 0.65 for the North, Center, West, East, and South regions respectively. Despite limitations such as the lack of data for validation outside France, and the non-standardized nomenclature for cover crops among Member States, this work constitutes the first effort to obtain a relevant cover crop map at a European scale for researchers and practitioners.

3.1 . Introduction

The Common Agricultural Policy (CAP), 2023-2027 introduces a series of reforms in key areas for the European Union (EU) Member States (MSs). These reforms aim to promote more intelligent, competitive, sustainable, and diversified agriculture and forestry, develop the socioeconomic structure of rural areas, and protect the environment. In particular, climate action, including carbon sequestration, is one of the EU's main priorities since it helps achieve the commitments made under the Paris Agreement. In this sense, conservation agriculture practices can help promote a shift of existing fields towards more sustainable systems. Such practices include minimizing soil disturbance, maintaining permanent ground cover, and adopting combined rotations (Hobbs et al., 2007; Palm et al., 2014). Among the existing options, the adoption of cover crops (CCs) constitutes one important example of a conservation measure to protect the soil surface against soil erosion. CCs are plants grown with the purpose of protecting the soil and improving its quality and health (Delgado et al., 2017). If CC biomass is incorporated into the soil, it positively impacts agronomic and environmental outcomes, such as soil carbon stocks (McClelland et al., 2021; McDaniel et al., 2014; Ruis and Blanco-Canqui, 2017). The adoption of CCs also has additional benefits, such as reducing nutrient leaching (Nyakatawa et al., 2001), improving soil structure through increased infiltration and water holding capacity (Nyakatawa et al., 2001; Panagos et al., 2015c; Smith et al., 1987), and improving the biological quality of the soil (Muhammad et al., 2021; Kim et al., 2020). CCs thereby play a strategic role in soil conservation policies since they are one of the few components of the erosion process that can be directly mitigated through human interventions by farmers and policy-makers (Panagos et al., 2015c). For their general contribution to soil health and climate change mitigation, CCs have been receiving increasing attention (Koudahe et al., 2022).

A key process justifying the implementation of CCs is soil erosion. Accelerated by human activities for more than 4,000 years (Jenny et al., 2019), soil erosion is currently the most common form of land degradation in the world, affecting over 1 billion ha of the Earth's surface (Borrelli et al., 2020). The average erosion rates worldwide are estimated to be 2.8-2.9 t $ha^{-1}yr^{-1}$ (Borrelli et al., 2017), around 3 to 20 times higher than the natural soil formation rates, of 0.15 - 1 t $ha^{-1}yr^{-1}$ (FAO, 2014). For the future, model forecasts indicate that current erosion rates may further increase from 30 to 66% worldwide (Borrelli et al., 2020), with local variations. In Europe, for example, the expected increase in soil erosion rates ranges from 13 to 22.5% (Panagos et al., 2021). As an established tool to combat erosion (Borrelli et al.,

2022), CC uptake has been shown to decrease erosion rates by 15 to 23% (Nyakatawa et al., 2001; Verstraeten et al., 2006) and mitigate on- and off-site damages (Montgomery, 2007), such as degradation of arable areas, pollution and eutrophication of rivers and lakes, and contamination of aquatic and marine ecosystems (Poesen, 2017; Amundson et al., 2015).

With the modern-day availability of Integrated Administration and Control System (IACS) data in the EU, another potential source of CC information is farmers' declaration, such as the Land Parcel Identification System (LPIS) and the Geospatial Aid Application (GSAA). These large-scale spatial databases contain annual declarations made for CAP measures by EU farmers, and their elements correspond to the boundaries of agricultural parcels and their corresponding main crops. While seemingly promising, these can be managed and operated independently at the regional or country level, so reporting secondary crops or CCs is sometimes possible but not consistent across MSs. Therefore, despite the increasing importance given to CCs, the availability of data about their use and presence remains primarily limited to coarse-scale statistical surveys. Among the European Union MSs, the best information available about CCs can be found in the Farm Structure Survey (FSS), which surveyed the 27 European countries down to the NUTS3 (the third level of the Nomenclature of Territorial Units for Statistics classification) regional level every 3 or 4 years (Commission, 2022a), and the Survey on Agricultural Production Methods carried out in 2010 (Commission, 2022e) at farm scale. While such information can be helpful for applications permitting spatial generalization, it does not provide sufficient spatial detail to allow, for example, a precise evaluation of the local impacts of CCs on soil erosion and carbon content.

Some consequences of lacking detailed and georeferenced CC data can be found throughout continental-scale modeling exercises. For example, when modeling the C-factor of the Universal Soil Loss Equation, Panagos et al. (2015c) used the CC area at NUTS2 level, the best available information at a European scale comprising 216 regions in the 27 MSs plus the United Kingdom (UK). In this case, the authors adopted the same value for all pixels within a NUTS2 region. Such an assumption is very strong and unlikely to represent field reality since it implicitly assumes that all detailed units (e.g., farms or parcels) behave identically. A similar assumption was also adopted in coupling erosion and carbon models to assess the combined effect of good agricultural and environmental practices on erosion and carbon budget at the national level (Borrelli et al., 2016). While the same assumption was not adopted by Borrelli and Panagos (2020), the lack of spatially detailed information about CCs was the limiting factor in their input datasets. In all these examples, a detailed CC map would

have represented an improvement in the characterization of physical processes.

While several initiatives to gather CC data exist, the information is not always publicly available, and no harmonized dataset is available at the EU level. This latter issue is difficult because the CC eligibility for CAP subsidies differs in each MS/region because of policy, management, and climate reasons. At the EU scale, the continent-wide Land Use/Cover Area frame statistical Survey Soil (LUCAS) field survey (Commission, 2022f; Orgiazzi et al., 2017) offers limited opportunity because it is predominantly made during the main cropping season and therefore cannot capture CC information. In addition to this, optical satellite imagery for the spatial monitoring of CC uptake often faces the challenge of a high incidence of clouds during rainy winter months in which CCs can be detected (Beriaux et al., 2021). The combination of these factors limits the availability of input data for automated computational techniques of CC detection with European-scalability. Even though the methodology for survey collections might change in the future, these reasons are probably why most recent computationally and data-intensive efforts using new technologies such as Copernicus Sentinel-1/2 satellites focus on mapping the main crop of the cropping season (d'Andrimont et al., 2021; Meroni et al., 2021).

Given the outlined scarcity of primary and auxiliary data, one possible alternative is using spatial disaggregation methods. Such a category of methods attempts to reconstruct the fine-resolution information from areal features (e.g. regional statistics polygons) to allow detailed spatial analysis (Comber and Zeng, 2019). With different approaches and assumptions, the application of disaggregation methods can be found in different disciplines, such as soil mapping (Møller et al., 2019), epidemiology (Utazi et al., 2018), disease mapping (Weiss et al., 2019), demography (Jia and Gaughan, 2016), among others.

This work focuses on the problem of disaggregating existing CC data in the European MSs and the UK from NUTS2 level to a finer spatial resolution using satellite data. The objectives are: a) to develop the first dataset of predicted CC occurrence at a high (i.e., 100-m) spatial resolution for Europe for 2016-2017; b) to develop the assumptions of a statistical model for the occurrence of CCs, which can be transposed to other study areas; c) to validate these newly produced maps quantitatively against parcel-scale observations; and d) to discuss some possible implications and applications of this new CC dataset. A statistical model for the occurrence of CCs is first built. The model uses 12-day median composites of remotely-sensed synthetic aperture radar (SAR) data from Sentinel-1. Its parameters are estimated iteratively in such a way that the aggregation of predicted values approximates the values reported by

statistical surveys at the coarse regional level. Then, parcel-level data from France's Registre Parcellaire Graphique (RPG) (Institut National de l'Information Géographique et Forestière, 2022), one of the few systems in European countries where declarations of CCs are public, is used to validate model predictions at the finest possible (i.e., parcel) spatial scale. Such an analysis allows understanding the predictions' strengths and limitations. Finally, some policy-relevant aspects of CCs are discussed.

3.2 . Methods

3.2.1 . Study area

The study area includes all croplands of the current EU MSs plus the UK, which covers about 156 million ha and 67 million parcels. Switzerland was not included due to the lack of harmonized survey data on cover crops. According to the Farm Field Survey (FSS), CC application has increased from 6.5% of all agricultural lands in 2010 to 8.9% in 2016 (Borrelli and Panagos, 2020). The adoption of CCs is currently an underused farming practice (Kathage et al., 2022) which is likely to increase in the EU in the future.

3.2.2 . Input data

Multi-temporal Sentinel-1 data was used to monitor changes in the landscape surface condition through time. The Sentinel-1 SAR constellation revisits the EU territory with a minimum 6-day revisit period since 2016 (until the Sentinel-1B defect in December 2021), providing a dense temporal time series for phenological monitoring. Compared to optical sensors (e.g. Sentinel-2), its microwave backscatter retrieval is practically uninfluenced by atmospheric conditions. In the agricultural context, microwave backscattering is sensitive to crop canopy structure, which means it can detect plant growth at the parcel spatial resolution. For these reasons, Sentinel-1 data offers a consistent source of plant phenological data in winter months for mapping CCs in combination with other computational techniques. Spatio-temporally consistent time-series of Sentinel-1 data were generated as follows: i) analysis-ready Sentinel-1 data was accessed in Google Earth Engine (Gorelick et al., 2017) (i.e., COPERNICUS/S1_GRD), already pre-processed to account for thermal-noise removal, radiometric calibration, and terrain correction; ii) a temporal stack of 31 rasters of the VV and VH bands was created, for both ascending and descending orbits, giving 12-day composites from the

10th of August 2016 to the same date in 2017; iii) the cross-polarization ratio (CR) was calculated as CR = VV/VH, giving 31 spatio-temporal layers for the period of study. Finally, the data was resampled using the median statistic to a 100-m spatial resolution, the target resolution adopted for the new spatially explicit CC dataset for reasons of computational trade-offs vs spatial resolution (data size) in spatial disaggregation models.

The choice of the CR time series as an indicator of the existence of CCs came from recent evidence highlighting the correlation of this index with the more well-known Normalized Difference Vegetation Index in the context of crop phenology (Meroni et al., 2021), crop dynamics (Veloso et al., 2017) and vegetation dynamics (Ma et al., 2022; Vreugdenhil et al., 2020). The choice of the period of analysis (i.e., 10th of August 2016-2017) was mainly driven by the need to match the period of most recent the coarse CC data available in Europe. In this case, it corresponds to the 2016 dataset on CCs per NUTS2 region published by the European Commission, which estimates the total CC area in arable lands (Commission, 2022b). The exact definition of the variable used (i.e., "cover or intermediate crops") reads:

> An area of arable land on which plants are sown specifically to reduce the loss of soil, nutrients and plant protection products during the winter or other periods when the land would otherwise be bare and susceptible to losses. The economic interest of these crops is low, and the main goal is soil and nutrient protection. (...) These crops should not be mistaken for normal winter crops or grassland (Commission, 2022c).

In order to filter the location of arable lands, the CORINE Land Cover 2018 at a 100-m spatial resolution and with a minimum mapping unit of 25 ha was used (Copernicus, 2022). By doing so, predictions of CC existence were restriced to a spatial domain where agricultural activity was previously detected in an external dataset. From the complete CORINE database, the class arable land was created by selecting the classes "non-irrigated arable land", "permanently irrigated land", and "rice fields". Additionally, other agricultural classes were included[1] to avoid having an overly restrictive spatial domain. Even though the period of CORINE does not match the period of 2016-2017, it can be accepted as adequate under the assumption that no drastic changes occurred in European arable lands from 2016

[1]Namely, "Vineyards", "Fruit trees and berry platforms", "Olive groves", "Annual crops associated with permanent crops", "Complex cultivation patterns", and "Land principally occupied by agriculture with significant areas of natural vegetation".

to 2017, the year where CORINE Land Cover 2018 images were mainly taken (Büttner and Kosztra, 2022). Then, because the amount of arable land calculated using FSS (Commission, 2022b) and CORINE are different, the CC area consistent with CORINE was defined by multiplying the original CC area from FSS by the ratio of CORINE to FSS arable land area (A), giving $CC_{Corine} = CC_{FSS} \cdot A_{Corine}/A_{FSS}$. Such an arable land ratio varies from 0.02 to 3.69, with an average of 1.02 (AP A.1).

3.2.3 . Disaggregation model

In order to disaggregate CC information from the (215) NUTS2 to the (156 million) pixels at 100-m resolution, the estimation method proposed by Fendrich et al. (2022b) was used. As in any other regression modeling approach, the technique consists of first making assumptions about the relationships between explanatory and dependent variables at the fine scale (Figure 3.1).

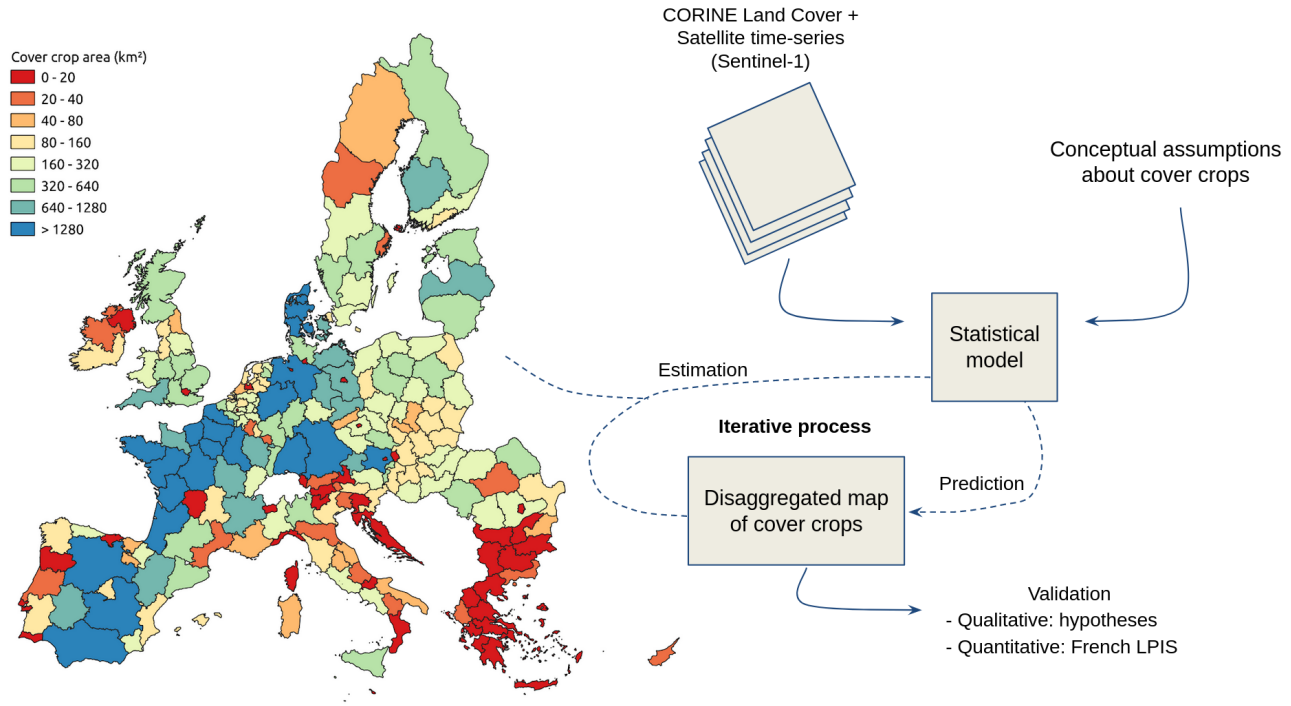


Figure 3.1: Summary of the method used in the present work. Left: area of cover crops in 2016 visualized using the cover crop extent per region from the Farm Structure Survey.

It was assumed that the fraction of CCs in a pixel varies in three dimensions, namely:

space, time, and the observed Sentinel-1 CR. Each of these dimensions are assumed to vary (possibly smoothly) with nonlinear interactions between them. Such an assumption is equivalent to assuming that the interpretation of a given CR time series varies so that it might indicate the existence of CCs in one particular place and its neighboring areas but not in other more distant regions. In mathematical notation, this can be represented as:

$$y_i = g\left[\sum_t s\left(lat_i, long_i, t, CR_i(t)\right)\right] + \epsilon, \quad \epsilon \sim \mathrm{N}(0, \sigma^2) \tag{3.1}$$

with y_i being the fraction from 0 to 100% of CCs within pixel indexed i; lat_i, $long_i$, t, $CR_i(t)$ being latitude, longitude, time and the CR time-series with a timestep of 12 days, respectively. The function $g(x) = 1/(1+exp(x))$ was chosen to be the link function, since y_i varies from 0 to 100%. The term ϵ is the residual term, assumed to be normally distributed, and $s(\cdot)$ is the joint function to be estimated from the data. Penalized smoothers that can be represented using basis expansions and a penalty matrix to control function smoothness are a common choice for the one-dimensional smoothers inside $s(\cdot)$, and the interaction is often represented as tensor products (Wood, 2017). Under such a representation, Equation 3.1 could be recognized as a nonlinear mixed model (Bates and Watts, 1988; Wolfinger, 1993).

In the representation of Equation 3.1, the fraction of CC at the pixel level is assumed to be a random variable. Consequently, the sum of the CC fractions in the pixels that belong to a NUTS2 region creates another random variable representing the CC area at the NUTS2 level. It can be shown that, in this case, a second nonlinear mixed model can be derived for the NUTS2 level, and it preserves the original parameters necessary to construct the pixel-level smoothers of Equation 3.1. Such a NUTS2-level model allows us to attempt to reconstruct the pixel-level information by performing regression analysis on the coarse data. During the parameter estimation phase, an optimization procedure handles the tradeoff between i) approximating the reconstructed to the observed values at the NUTS2 level; and ii) enforcing the mathematical assumptions for the function $s(\cdot)$ described above (i.e., that the one-dimensional smoothers are continuous functions, that the effect of the CR varies in space and time etc.).

Parameter estimation is done through a numerical optimization procedure, which maximizes the chance of observing the aggregated data given the assumptions made at the pixel level (i.e., maximum likelihood estimation). In this work, we slightly modified the original method of Fendrich et al. (2022b) to use a quadratic instead of a first-order approximation

to the log-likelihood in each iteration. Such a modification led to the natural interpretation of each iteration as a Newton-Raphson step of the original guess toward the maximum a posteriori estimates, similar to the scheme proposed by Rossell et al. (2021).

As disaggregation problems are fundamentally undetermined with infinite solutions, choosing proper assumptions is essential for narrowing the possibilities. Therefore, $s(\cdot)$ was assumed to be a tensor product, and several possibilities for the one-dimensional smoothers of Equation 3.1 were tested, including P-, cubic and thin-plate splines, with different basis dimensions. AP A.4 shows the results for the four best models found during model selection according to the area under the curve (AUC) performance metric (see subsection 3.2.4 for details about the validation dataset). The final alternative was chosen to be four P-splines (Wood, 2016), with basis dimensions 9 for the longitude and 8 for latitude (AP A.3), time and CR. After the estimation of model parameters, the expected value of $s(\cdot)$ given the parameters was calculated and plotted to allow the visualization of the estimated one-dimensional smoothers. The predicted CC fraction and the corresponding uncertainty were calculated as the point estimate of the median and the bounds of the 90% confidence intervals for the expected value of the response variable of Equation 3.1 given the parameters, respectively.

3.2.4 . Validation

Qualitative validation

Validating a disaggregation model is a challenging task. First and foremost, no fine-scale information is available for the whole study area. Otherwise, if such data were available for comparison against model predictions, it would be preferable not to use a disaggregation model but instead to use such information as an input to traditional mapping techniques. Given these considerations, the validation procedure was split into: i) a qualitative validation of the model's internal behavior; and ii) a quantitative validation of the external behavior (classification accuracy), consisting of a comparison against publicly available parcel data for France.

Developing a reasonable interpretation of the results obtained from a prediction is a vital step in regression modeling (Nisbet et al., 2009). Therefore, in the qualitative validation, the estimated smoother (Equation 3.1) was investigated to check its representation of logical aspects of CC phenology that should reasonably result in their identification. To do so, necessary but not sufficient conditions were first established for the model to be realistic. The

first condition is for the model to assign a higher weight to the satellite observations made during cold months compared to hot months since that is when CCs can be observed in the field. The second condition is for the model to capture the different regional patterns of CC production properly based on the known a priori information. This secondary condition checked, for example, if the model could capture the regionalized temporal pattern of CCs in France, where sowing dates may vary from the end of July in the Moselle region (North-West France) until the beginning of November in the Pyrénées-Atlantiques (South-West France) (Journal Officiel de la République Française, 2018). In order to identify this information from Equation 3.1, a first-order expansion of g(x) around $x = 0$ was used to calculate the approximate marginal contribution of the CR to the predicted fraction of CCs at the pixel level. Then, since the model of Equation 3.1 is an interaction between several variables, only the marginal results were presented.

Quantitative validation

To complement the qualitative evaluation of the model, a quantitative validation was also developed to confront model predictions against comparable parcel-scale declarations made by farmers in the CAP context. Such a validation started by searching for the most accurate classification of parcels between those with or without CCs in 2016. To do so, databases of European countries (Schneider et al., 2021) were inspected, and it was found that in most cases, farmers are only required to report their parcel's main crop. Nevertheless, in some countries, such as Portugal and France, farmers can, but are not required to, report multiple crops per year.

For the present work, the French RPG dataset was chosen, which gathers all annual declarations made by farmers to receive CAP subsidies (Levavasseur et al., 2016) and provides the cultivated crop or crop groups for each farmer at the parcel scale. This private and confidential information concerns the parcels belonging to farmers provided by the French Ministry of Environment. The RPG dataset was used to validate the predictions for three reasons. First, France is representative for being the largest MS, with a large share of its territory devoted to agricultural activity, and for covering several environmental conditions such as continental, Mediterranean and oceanic climates (Ols et al., 2020). Second, as shown in Figure 3.1, France contains a large area of CCs, and the inclusion of CCs in the cropping system is, in some cases, mandatory to comply with the Nitrate directive (of the European Union, 1991) and the CAP Greening (J. Kathage, I Pérez Domínguez, 2019). Third the French Ministry of Agriculture and

Food Sovereignty provides a list of the 44 crops that are considered to be CCs (Ministère de l'Agriculture te de l'Alimentation, 2022) to comply with the dedicated Ecological Focus Areas (EFAs) to safeguard and improve biodiversity, requested by the green direct payment of the CAP. Such a list facilitates farmers' declarations, which have to declare two of these crops in a mixture per parcel to be retained as EFAs.

However, the construction of the validation dataset faces some challenges. As the dataset is attached to compliance with regulations, motivations to meet the minimum requirements in the CAP can bias the area of CCs declared by farmers. For example, a farmer with more CCs than the minimum requirements may feel inclined to declare only the minimum, which would generate a bias towards underestimating CCs. Conversely, a farmer with half the minimum requirement could tend to overestimate the area of CCs in the declaration. Besides, the declaration of CCs is not mandatory in France due to other EFAs that can be present on the farm. Therefore, it is not possible to affirm with certainty whether a random parcel contains CCs or not. This problem was overcome by applying additional filters on the dataset to increase the confidence in the presence or absence of CCs on the parcels. Such a procedure aimed at obtaining a sample of parcels that could be used with greater certainty to validate the model at the scale of reference using the available data. The filters are described next.

First, three groups of parcels according to the different levels of uncertainty of CC existence were defined a priori: parcels without CCs in farms that did not declare CCs (PnFn, lower certainty); parcels without CCs in farms that declared CCs (PnFy, medium certainty); parcels with CCs in farms that declared CCs (PyFy, higher certainty). The sampling procedure for the PnFn group consisted of first filtering farms without CCs declared, and then randomly sampling parcels inside these farms. This group has the lowest certainty among the three because as no CCs are declared within the farms, the farmers may have chosen to declare other EFAs in their farms instead of CCs. In this case, some farms might have CCs on the field but not registered in the database.

For PnFy and PyFy, two filters were added at the farm level: i) only those with a sufficiently high ratio of surface as CCs on the farm (i.e., 20%), compared to the minimum cover limit imposed by the CAP (5%); and ii) only those with a large area (over 50 ha) to increase the chance of having a heterogeneous configuration of parcels. These filters assume that if a farmer has declared well beyond the required minimum of 5%, then he has declared the entire area of CCs in his parcels. However, since this assumption may not be true in some cases, more confidence can be assumed for the presence (PyFy) than the absence (PnFy) of

CCs in this case, which justify their certainty assigned above. Finally, parcels whose polygons fall entirely inside the CORINE land cover classes described in subsection 3.2.2 were randomly sampled for each of the three classes.

After these procedures, 19,390 parcels were used for validation: 8,441 PnFn, 11,112 PnFy, and 12,256 PyFy, each with information about CC's existence and its main crops in 2016. Among the 19,553 sampled parcels without CCs (i.e., PnFn + PnFy), 5,968 had a winter commercial crop as the main crop. Finally, because the size and shape of each land parcel are different, they were compared by calculating the weighted median value of all model predictions intersecting their boundaries. Boxplots were plotted to compare the distribution of predictions within the three classes of parcels described above, and receiver operating characteristic curves (i.e., ROC curves) were generated to assess the model's discrimination abilities.

For each ROC curve, the corresponding AUC was calculated. Their interpretation was made using the classification proposed by Bera et al. (2020). Weak: $0.5 \leq$ AUC < 0.6; Moderate: $0.6 \leq$ AUC < 0.7; Good : $0.7 \leq$ AUC < 0.8; Very good: $0.8 \leq$ AUC < 0.9; Perfect: AUC ≥ 0.9. Analogous quantitative-qualitative relationships can be found in works such as Roy et al. (2020); Panahi et al. (2022), and others.

3.2.5 . Implementation

All codes used were written in R 4.2.0 (R Core Team, 2022). The P-splines implementation used was that of the mgcv package (Wood, 2017), spatial data was manipulated using the terra package (Hijmans, 2022), and parallel computations were made with the snow package (Tierney et al., 2021). The row-wise Kronecker product and the log-likelihood were implemented with Rcpp and RcppArmadillo (Eddelbuettel and Balamuta, 2018; Eddelbuettel and Sanderson, 2014). All source codes and input datasets used, except for the parcels' boundaries used for validation, are publicly available in a dedicated GitHub repository.

3.3 . Results and discussion

3.3.1 . A qualitative evaluation of the model response

For the marginal spatial effect, Figure 3.2 shows the results for a winter month, December 2016. Six different locations are shown to highlight that each location has a distinct pattern for the effect of CR on the calculated fraction of CCs. For example, in Figure 3.2, the model curves for South Italy and Bulgaria show that the low CC fractions (expected in these regions according to Figure 3.1, left) are represented through a negative effect of the CC fraction for practically the full CR range. Conversely, for the other four locations where CCs tend to be more widespread, according to Figure 3.1 (left), a different pattern can be observed. In these cases, Figure 3.2 shows that an increase in the CC fraction is predicted for most of the CR range, and regional variability in the shape of the curve seems to be captured. As a reference for the CR values, Meroni et al. (2021) suggests a range from 0.1-0.2 in the trough (minimum) to 0.4-0.5 in the peak (maximum) of temporal series of crop fields in Europe (i.e., common wheat, rape, and maize), approximately the same range reported by Maurya et al. (2022) and Vavlas et al. (2020) in wheat fields in the North of India and the UK, respectively.

A distinct pattern in the marginal spatial effect can be seen in North-West France, where the effect on the CC fraction quickly peaks around 0.20 and decreases, and in Denmark, where the effect has a double peak, with the second starting around 0.20 and reaching its maximum around 0.30 (Figure 3.2). In both cases, the fact that the model assigns the highest increase in CC fractions to relatively low CR values could mean that it is attempting to distinguish CCs from commercial winter crops, which could present a different signal due to fertilizer application. Such an interpretation is however uncertain since it could also represent other factors that affect the CR value, such as soil surface roughness from tillage operations (Vreugdenhil et al., 2020).

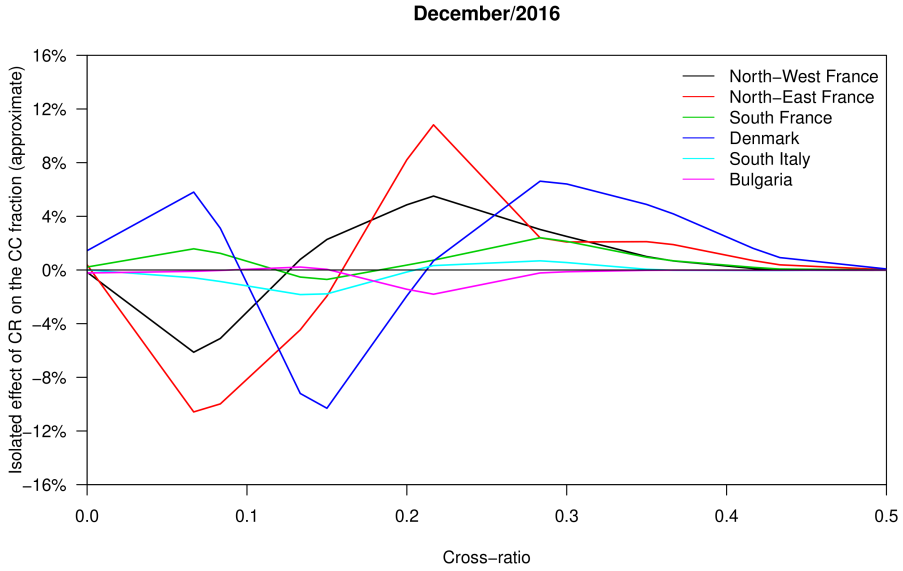


Figure 3.2: Qualitative validation, marginal spatial effect: (approximate) isolated effect of the Cross-ratio (CR) on the predicted cover crop (CC) fraction as estimated by the model, for one of the 12-day median CR values in December/2016.

For the marginal temporal effect, the results of Figure 3.3 are presented for Denmark (i.e., the blue line of Figure 3.2). The Figure shows that, when weighing between available observations, the model attributes a relevant effect from the observations made in all months of the year, which can be seen by the large overall positive and negative effect on the CC fraction. For example, for low CR values ranging from 0.1 to 0.2, spring months have a comparable effect to that of winter but in the opposite direction. Such a result resembles the observations made by Nowak et al. (2021), who showed that different spring-grown crops in France present different soil cover patterns during cold months. The hypothesis that cold months could have a higher influence than hot months seems to be valid only for CR values above 0.2. In practice, these results suggest that the approach to detect winter soil cover based on satellite data for winter months only proposed by Nowak et al. (2021) might be insufficient for CCs. It also suggests that a complete time series must be considered to separate CCs from commercial winter crops. In the latter case, the combination of complete time series and a flexible and probabilistic model could better identify what is a CC within a crop succession in a context-dependent way.

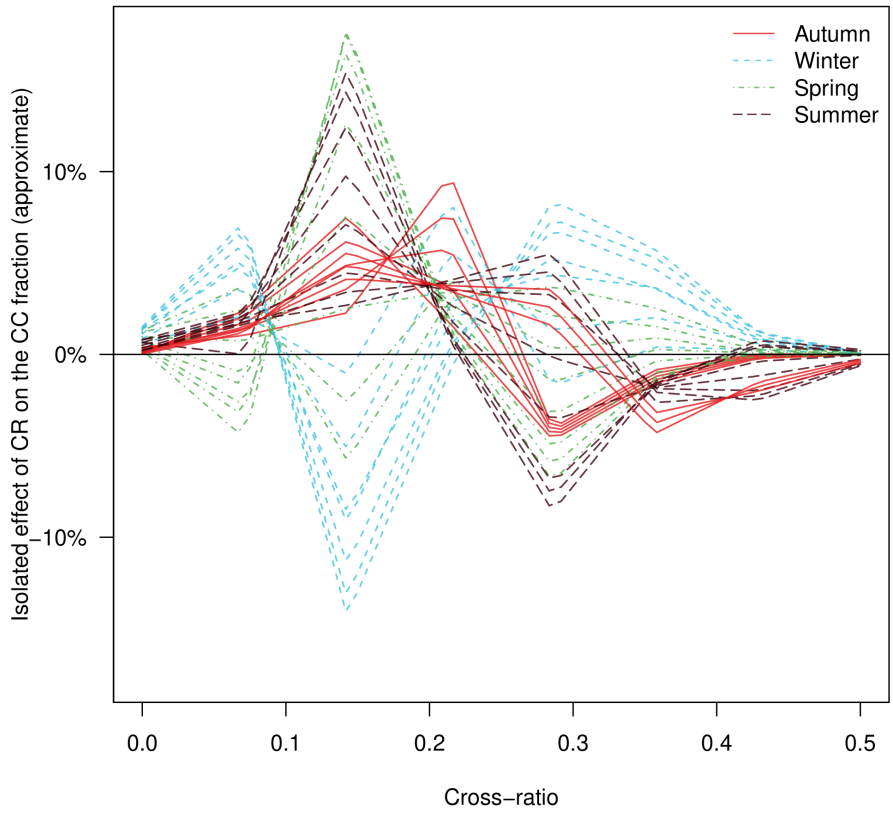


Figure 3.3: Qualitative validation, marginal temporal effect: (approximate) isolated effect of the Cross-ratio (CR) on the predicted cover crop (CC) fraction as estimated by the model, for a random location in Denmark. Every line corresponds to one of the 12-day median CR values used for fitting the model.

3.3.2 . A quantitative evaluation of model performance

The calculation of weighted median predictions for all parcels (i.e., n >= 1 pixel) and only parcels intersecting more than 10 pixels (i.e., n >= 10 pixels) is shown in Figure 3.4. For the case of all parcels included, the model could predict higher values for PyFy (right) than PnFn and PnFy (left, center), which indicates an ability to distinguish between parcels with and without CCs, respectively. The proximity between the distributions of PyFy and PnFy seems to indicate some confusion between parcels of farms known to have CCs. Such a result could indicate that the model performs better when the adoption of CCs is widespread at the farm level. However, such confusion seems to be minimized when only parcels intersecting more than 10 pixels are included in the analysis. In this case, PnFy approaches PnFn, with lower values predicted than PyFy.

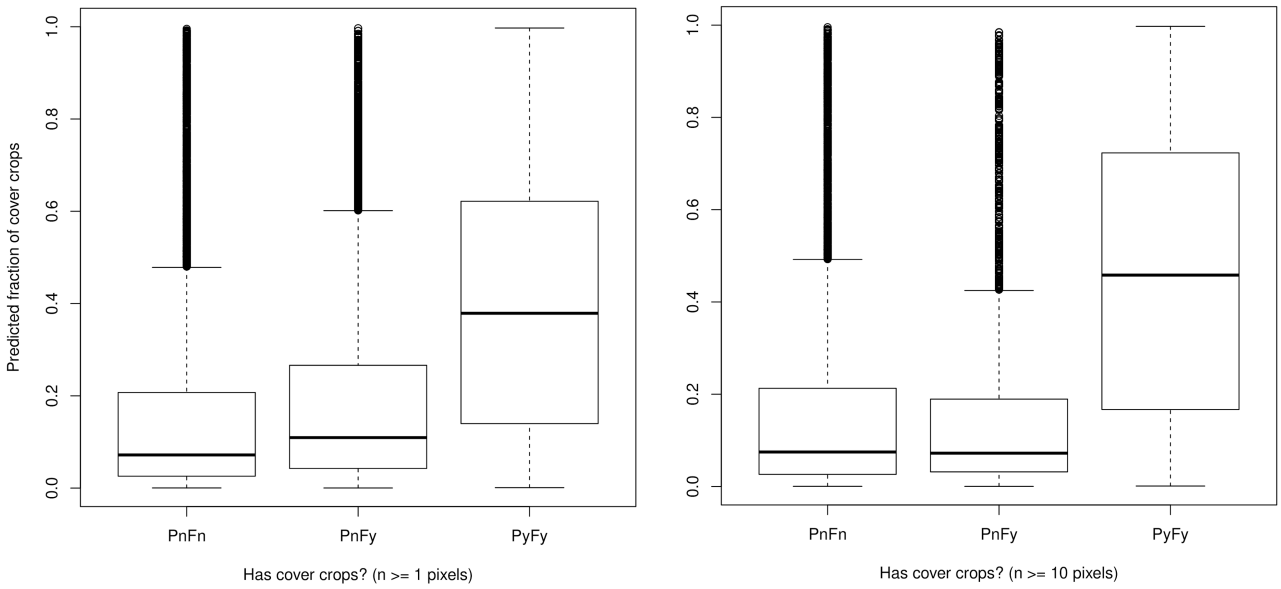


Figure 3.4: Quantitative validation: distribution of the weighted median values calculated for French parcels in arable lands. Two cases: all parcels included (i.e., n >= 1 pixel, mean area of 5.2 ha) (left), and only parcels intersecting more than 10 pixels included (i.e., n >= 10 pixels, mean area of 9.6 ha) (right). PnFn are parcels without cover crops in farms that did not declare cover crops; PnFy are parcels without cover crops in farms that declared cover crops; PyFy are parcels with cover crops in farms that declared cover crops. The separation between n >= 1 and n >= 10 was made to evaluate the impact of sub-pixel confusion (a mix of landscape dynamics at the sub-pixel level) in the input data

Despite the results shown in Figure 3.4, a question could be posed about the model's ability to distinguish between different types of winter cover. To evaluate this matter, Figure 3.5 shows the analysis results that use crop information at the parcel level to separate winter commercial crops from CCs. Parcels with winter commercial crops were selected based on the sowing period of the main crop reported to the RPG, excluding CC species. It can be seen that a similar pattern as Figure 3.4 is found, indicating a higher predicted fraction of CCs in parcels that truly contain CCs. Such a result is relevant, as it indicates that the assumption hypothesized in Equation 3.1 of a nonlinear effect of CR on the CC fraction is adequate to differentiate between the multiple signal patterns of soil cover in winter.

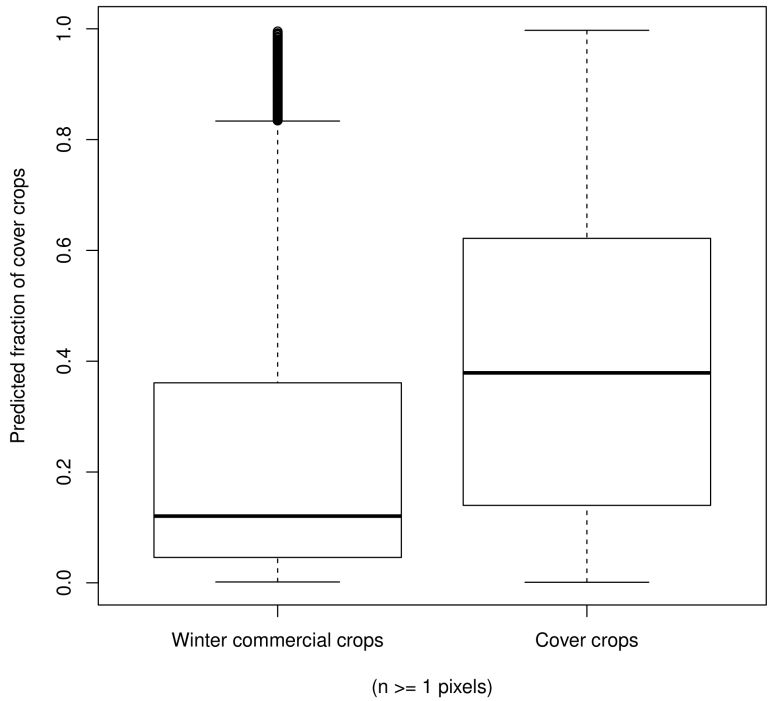


Figure 3.5: Quantitative validation: distribution of the weighted median values calculated for French parcels with winter commercial crops and cover crops. The left boxplot merges all parcels with winter commercial crops within the groups PnFn (parcels without cover crops in farms that did not declare cover crops) and PnFy (parcels without cover crops in farms that declared cover crops).

While Figure 3.4 and Figure 3.5 show the general pattern of model validation in France, regional patterns to understand the model's biases were also investigated. Figure 3.6 shows the reclassification of the French NUTS2 areas into five contiguous regions (namely, South, East, North, West, and Center), and the model validation within these five areas. For each region, the ROC curve shows the true positive rate in the vertical axis against the false positive rate in the horizontal axis, summarizing the model performance when used as a binary classifier for the presence/absence of CCs in French parcels. The ROC curve is presented here, along with its corresponding AUC (Maimon and Rokach, 2010). In all regions and the whole of France, the AUC calculated is greater than 0.5, indicating that the model performs better than random chance. The results also show that, compared to the overall performance of AUC = 0.74, the West region of France yields an identical AUC, the North and Center regions yield a higher AUC (0.77 and 0.75, respectively), and the South and East regions a lower AUC (0.70 and 0.65, respectively). According to the classification presented in subsection 3.2.4,

this means that the overall classification and 4 out of 5 regions presented a good performance, while the East region had a moderate performance. The variable pattern across the country also shows that regional classification biases exist in the model and that more errors tend to happen in the South and East regions of France.

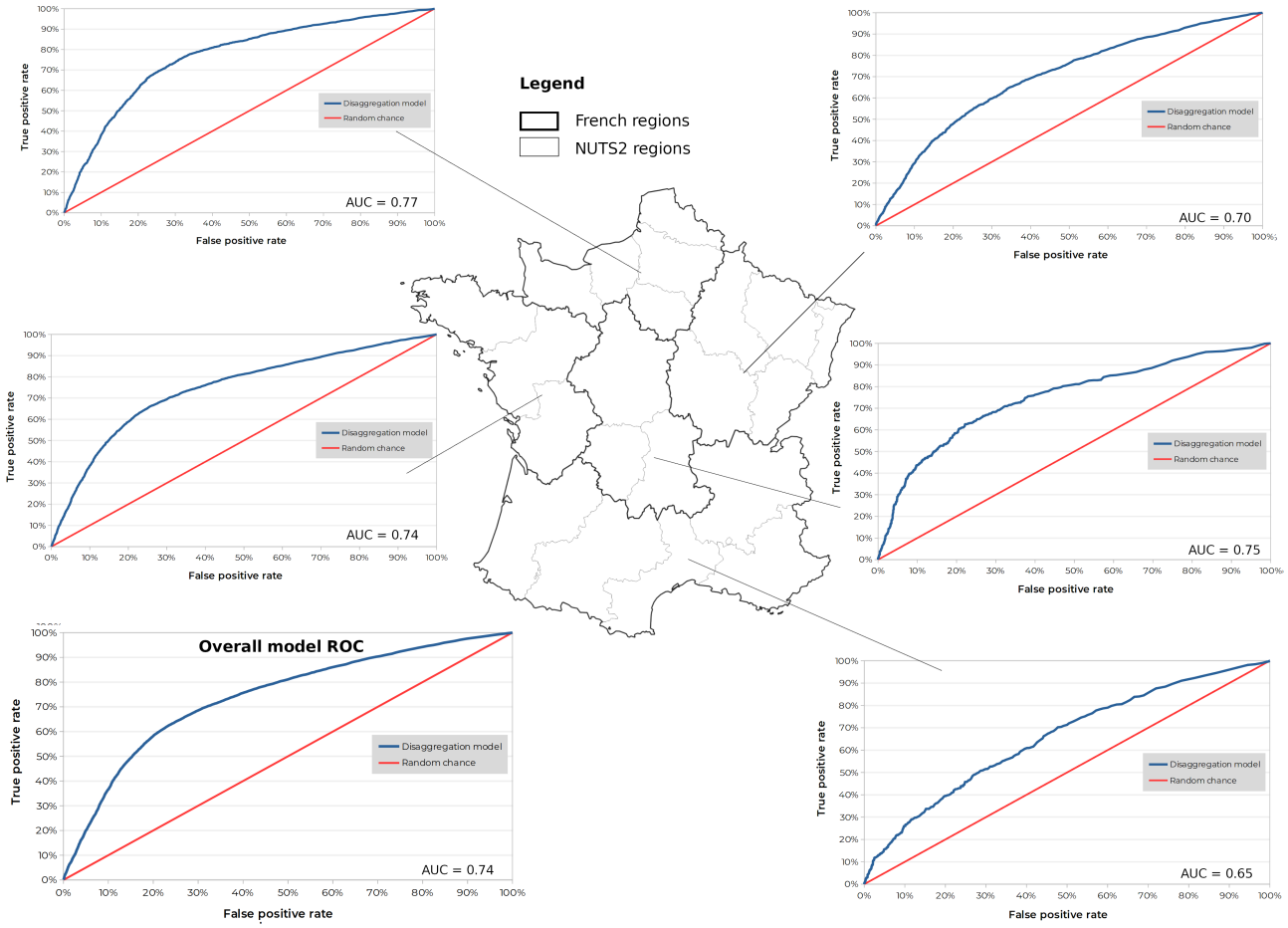


Figure 3.6: Quantitative validation: regional variability of model predictions across France.

For the European-wide model (Figure 3.6, bottom left), a threshold value was calculated as being the point in the ROC curve at which the Euclidean distance to the theoretical optimum is minimized (i.e., the top-left corner). The confusion matrix generated for this threshold value (Table 3.1) shows that when parcels contained CCs in the RPG, the model predictions were correct in 68.1% of the cases. When the parcels did not contain CCs, the model was correct 74.6% and 67.5% for PnFn and PnFy, respectively. It also shows that, when the model predicted the inexistence of CCs, it was wrong in 22.1% of the cases, and therefore correct

in 77.9% of them. When the model predicted the existence of CCs, it was correct in 59.2% of the cases and thus wrong in 40.8% of them. In this case, this model seems to assign more false positives to PnFy (25.6%) than to PnFn (15.2%), which can be possibly explained by some farmers not declaring all their CCs due to other EFAs in the farm (see subsection 3.2.4). While these results suggest a limited power of the model in differentiating CC presence, it is necessary to highlight that: i) they implicitly assume a single threshold value for the whole of France, which can be suboptimal according to the heterogeneous results presented in Figure 3.6, and ii) another threshold to improve model's precision could be chosen. In this context, the use of continuous values (e.g., Figure 3.7) might be appropriate depending on the application intended.

	Model prediction : contains CCs?	
Situation	**No**	**Yes**
PnFn	6,296 (74.6%)* (35.6%)**	2,145 (25.4%)* (15.2%)**
PnFy	7,498 (67.5%)* (42.4%)**	3,614 (32.5%)* (25.6%)**
PyFy	3,907 (31.9%)* (22.1%)**	8,349 (68.1%)* (59.2%)**

* refers to the 'row' percentage
** refers to the 'column' percentage

Table 3.1: Quantitative validation: confusion matrix for the overall model in France. The results are based on the assumption of a single threshold for the whole country.

Figure 3.7 shows the model predictions for the arable land in the whole of Europe. The continental level map shows that the fraction predicted in most pixels with arable land tends to be zero and that CCs occur in concentrated regions but with local variations (Figure 3.7, zoom). The marginal distributions show that CC fractions tend to be higher in the regions corresponding to Northern France, Germany, and Denmark, following Figure 3.1. The standard deviation of model predictions, as well as the 0.05 and 0.95 quantiles are provided in the Appendix (AP A.5 and AP A.6). These uncertainty margins were generated using the uncertainty around the estimated smoothers.

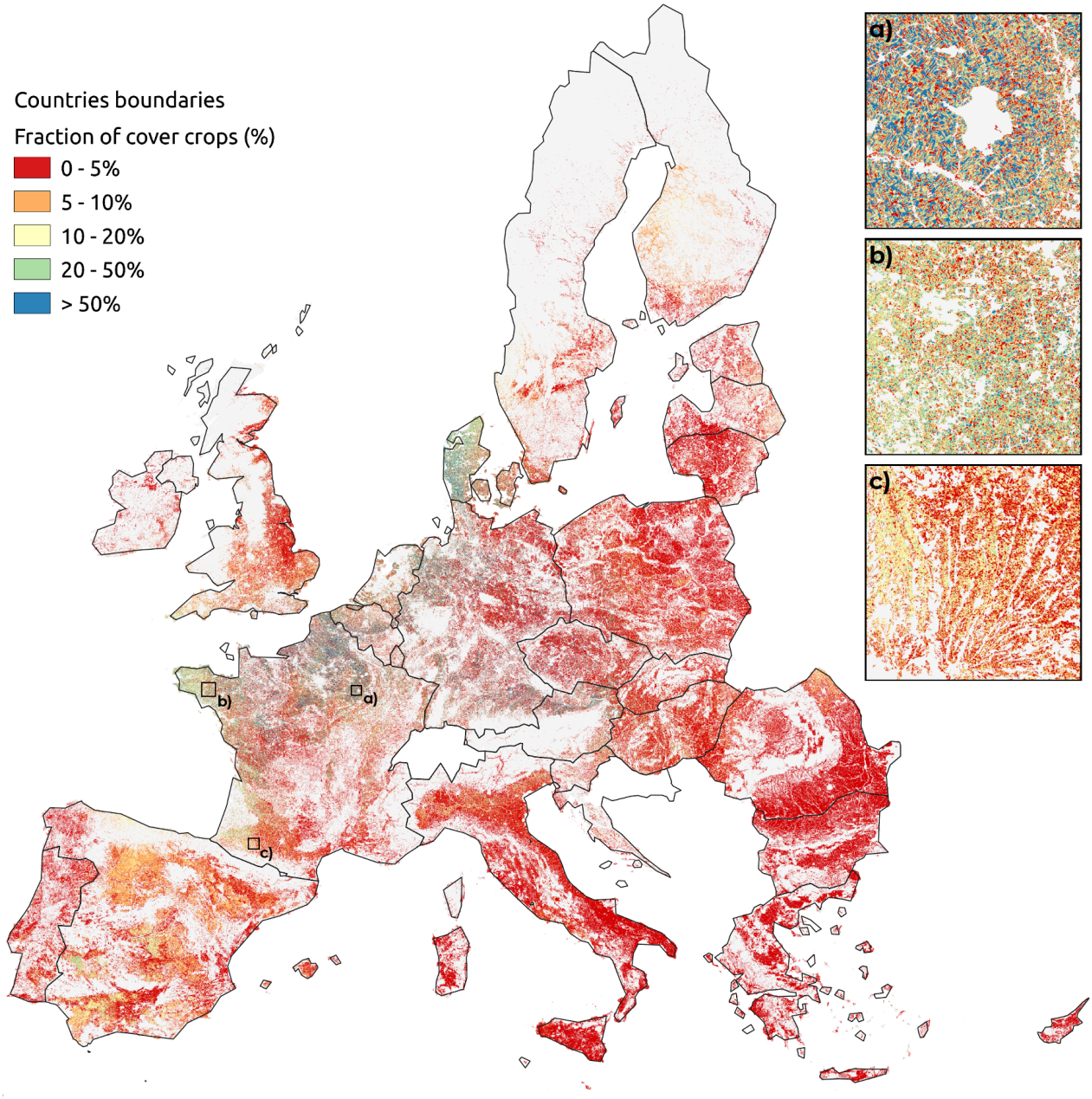


Figure 3.7: Left: model predictions of the occurrence of cover crops (CCs) in Europe. Right: Three zooms: Predictions on the East, West and South of France (a, b and c, respectively).

3.3.3 . Future policy opportunities offered by spatially explicit cover crop predictions

The map produced in this work is the first European product at a 100-m spatial resolution predicting the fraction of CCs for the winter season of 2016-2017. Despite the limitations,

with a lack of validation outside France and the existence of mixed satellite signals (subsection 3.3.4), our dataset fills a gap and can be tested in several future applications ranging from fine-resolution input data for analysis of soil loss by water erosion, organic carbon sequestration, nutrient application, and to assist the implementation of soil conservation policies at a regional or national scale. Given that the FSS survey data for 2023/24 is still under collection at the time of writing, the current work also provides a reference methodology to be used for the spatial disaggregation of CCs and other policy-relevant information (e.g. reduced tillage). Such a temporal analysis would enable the possibility of evaluating trends in CC adoption in recent years.

Two main policies are currently responsible for driving European farmers' adoption of CCs; the Nitrates Directive, and the CAP. Such policies, especially the Nitrates Directive, are now the strongest determinants of the adoption rates and intensities of CCs by farmers (Kathage et al., 2022). Such factors overcome agronomy or environmental motivations, although this scenario might change in the future as CCs become economically incentivized, for example, for energy production (Launay et al., 2022). The spatial pattern of the vulnerable zones under the Nitrate Directive in France resembles the areas with high fractions of CCs in the map produced by this study (Figure 3.7) (JRC, 2022). Furthermore, covering the soil in the most sensitive periods, a practice that includes adopting CCs, constitutes one of the Good Agricultural and Environmental Conditions (GAECs) introduced in the upcoming CAP 2023-2027 to protect soils in rainy susceptible seasons (e.g winter). In the proposal, the European Commission introduced various scenarios to protect soils using soil conservation measures such as the minimum green cover. The adoption of CCs with a cover rate of 75% in one of the policy scenarios led to a reduction of soil erosion in arable lands and permanent crops by 15 and 30%, respectively (Commission, 2018; Panagos et al., 2021). Therefore, the new CAP 2023-2027 is likely to increase the area of agricultural land with CCs in the EU.

One of the GAECs of the new CAP prescribes that MSs must require farmers to apply crop rotation (Commission, 2022g), a practice that also composes the list that can be supported by eco-schemes (Commission, 2021), and a summary of the first strategic plans shows that 16 MSs include the practice of crop rotation with a secondary or CCs during the winter season to protect soils (Commission, 2022g). In this sense, while the currently available CC data at the NUTS2 level collected by the FSS is useful to monitor the overall adoption rates and intensities of CCs, these coarse data are less suitable for evaluating the effect agricultural practices by farmers on the environmental and climatic objectives that are set by European

policies (Matthews et al., 2023). By contrast, the use of techniques to develop refined spatial datasets as produced in this study is highly valuable to assess the influence of European policies on the on-ground-farmers decisions and the subsequent impact on climate and the environment. Since the new CAP emphasizes performance and results, assessments with detailed data like the one produced in this study will become more important.

3.3.4 . Model limitations and future improvements

While the present results represent an advance for CC mapping, several limitations of the current approach can be listed and possibly developed in further works. The disaggregation model is strongly dependent on the FSS CC data at the NUTS2 level. In the last decades and once every ten years, the FSS has been a part of the Census of Agriculture, which guarantees high reliability of the values provided. However, the values available for 2016 and used in the current work correspond to sample estimates (Commission, 2022d). This means that not only is the sample size used relatively small, but also the data itself contains uncertainties. Although the disaggregation model mitigates the problem by not enforcing the equality constraint but only approximating (Fendrich et al., 2022b) (AP A.2), more reliable data could potentially lead to more precise results. In the near future, area monitoring systems have the potential to provide more information about CCs at the EU level (of the European Union, 2021).

Here, the spatial resolution of 100-m was adopted, approximately ten times the original resolution of the Sentinel-1 data. In practice, this means that the pixels in the present work can potentially have mixed signals in areas of high heterogeneity, which adds an extra layer of uncertainty where non-unique cultivations are found in close proximity. As shown in Figure 3.4, this effect directly impacts the predictions for smaller parcels but tends to be minimized for larger units. Even though the model could be run at 10-m resolution, this would bring new computational and practical challenges since the number of pixels to predict would increase 100-fold. Such a high-resolution application would also require highly accurate land cover masks to be used as spatial input data. While the CORINE croplands dataset, as used in this study, is widespread in academic studies, it does not capture the fine-scale variability at the parcel level. Replacing it in the future can be a direction for further work.

The added additional benefit of a standardized validation dataset at a continental scale would be very relevant, since management practices present a high variability across the EU

(Panagos et al., 2015c). Due to the current lack of information, it was necessary to assume in this work that if the model presents reasonable results for France, then it is appropriate for other regions. France is a particularly good choice of MS due to data availability, the current status of CC implementation, and its high diversity of climate regions. However, the fact that the performance of the results in other regions is still unknown and must be improved in further work. Therefore, we highlight the need for increased data availability and an accompanying definition of what does or does not belong to the CCs group among EU countries. In this sense, it is necessary to reinforce that the results obtained in the present work are conditional on the broad definition presented in subsection 3.2.2. As an outcome of this work, we justify the need for a move towards a consistent terminology in the EU in order to facilitate the synergistic use of datasets and improve future prediction exercises.

3.4 . Data availability

The datasets of predicted CC fraction, standard deviation, and the 5% and 95% quantiles at 100-m spatial resolution are available in the European Soil Data Centre (ESDAC) (JRC, 2023; Panagos et al., 2022a).

3.5 . Conclusions

In this work, a statistical disaggregation model is proposed to derive CC information at 100-m resolution from aggregated statistics reported in the Farm Field Survey at the NUTS2 level in Europe (i.e., 215 regions with CC information). The transference from the coarse (regional aggregations) to the fine (pixel) level was made through a statistical model constructed based on assumptions relating CC phenology in arable lands to a full annual time series of the cross ratio (CR) from Sentinel-1. To then quantify the model accuracy, the best available data sampled from spatially-explicit farmers' declarations in the French RPG was used to validate the model at the field parcel scale. The models' interior behavior was shown to be coherent with reality, modifying its sensitivity to the Sentinel-1 CR time series according to the region and moment of the year under consideration. In regions with known CC implementation, the model properly identified the importance of considering a complete time series to generate context-dependent model predictions. This multi-temporal approach was important in the model's successful distinction of different types of winter cover, in which lower fractions of CCs were predicted in places with winter commercial crops. The results showed that in general the model was able to successfully predict higher CC fractions in areas where they are planted in fields. Overall, when interpreted as a binary classifier for the whole of France, the model yielded an Area Under the Curve (AUC) of 0.74. On a regional basis, the AUC values were 0.77, 0.75, 0.74, 0.70, and 0.65 for the North, Center, West, East, and South regions, respectively, showing geographical variation in the model accuracy. Despite discussed limitations, this derived data layer can provide an important and updateable information source for researchers and practitioners requiring a spatially explicit knowledge of CC implementation.

3.6 . Learnings

In this chapter, we investigated the possibility of generating spatially explicit cover crop information using a disaggregation model to combine fine-scale satellite data and coarse-scale statistics about cover crops. The results pointed to the following learnings:

- Despite the data scarcity setting, the cross-ratio (CR) signal from Sentinel 1 is a potential good source to indicate cover crop existence in Europe, with evidence suggesting an ability to separate cover crops even from winter cash crops.
- The assumption that the CR signal indicating cover crop existence varies in space seemed valid, given the different curve patterns obtained for different countries sampled in Europe.
- The variable model performance across regions in France indicates that other variables could help refine the results, and further work is necessary to investigate this possibility.
- The lack of publicly available parcel data precludes research on cover crops. Even though cover crops are a particularly tricky subject to work with at a large scale due to the lack of information from optical satellites in winter months, parcel data availability could be an essential source of information for researchers.

Despite the limitations of validating the results only for France and not the other European countries, this chapter produced the first map of cover crops at a continental scale. We hope it will be helpful for other researchers to improve assumptions that are often made when modeling cover crops. In this thesis, combined with the model developments made during this thesis, the new map allowed us to simulate policy-related scenarios in Chapter 4.

4 - Improving land management representation in a constrained European carbon scheme with lateral displacement to oceans

This chapter was written by Arthur Nicolaus Fendrich, with contributions from Philippe Ciais, Panos Panagos, Philippe Martin, Marco Carozzi, Bertrand Guenet, and Emanuele Lugato.

The use of cover crops (CCs) is a promising cropland management practice with multiple benefits, notably in reducing soil erosion and increasing soil organic carbon (SOC) storage. However, the current ability to represent these factors in land surface models remains limited to small scales or simplified and lumped approaches due to the lack of a sediment-carbon erosion displacement scheme. This precludes a thorough understanding of the consequences of introducing a CC into agricultural systems. In this work, this problem was addressed in two steps with the spatially distributed CE-DYNAM model. First, the historical effect of soil erosion, transport, and deposition on the soil carbon budget at a continental scale in Europe was characterized since the early industrial era, using reconstructed climate and land use forcings. Then, the impact of two distinct policy-oriented scenarios for the introduction of CCs were evaluated, covering the European cropping systems where surface erosion rates or nitrate susceptibility are critical. The evaluation focused on the increase in SOC storage and the export of particulate organic carbon (POC) to the oceans, compiling a continental-scale carbon budget. The results indicated that CCs simultaneously increase SOC storage while reducing POC export to the oceans. Compared to the simulation without CCs, the additional rate of SOC storage induced by CCs peaked after 10 years of their adoption, followed by a decrease, and the cumulative POC export reduction stabilized after around 13 years. The findings indicate that the impacts of CCs on SOC and reduced POC export are persistent regardless of their spatial allocation adopted in the scenarios. Despite some known limitations, the current work constitutes the first approach to successfully couple a distributed routing scheme of eroded carbon to a land carbon model emulator at a reasonably high resolution and continental scale.

4.1 . Introduction

Among the many management options for a more sustainable cultivation, the adoption of cover crops (CCs) has recently received increasing attention for its potential benefits (Scavo et al., 2022). CCs are grown during the fallow period and between two successive main crops, interrupting their cycle before competing with the next main crop. This practice, often associated with reduced tillage techniques, improves soil fertility through root exudates and the return of litter and biomass to the soil. Similar actions are also referred to as 'green manure' or 'catch crops' when CCs are introduced specifically to add carbon and nitrogen, or to retain nitrogen by reducing leaching, respectively (Shackelford et al., 2019). In 2009-2015, the United States of America's Environmental Quality Incentives Program reported CCs as one of the most common conservation agriculture practices, with an incentive funding that increased from US$ 15 million to US$ 56 million per year (U.S.Government Acocuntability Office, 2017). In 2015, the European Union (EU) introduced the adoption of CCs in the Common Agricultural Policy (CAP) as an option for the Ecological Focus Areas (EFAs), requiring farms with more than 15 ha to devote at least 5% of their arable land to environmental and climate-related activities, including the use of cover crops. In its first year of application, 27.7% of the land devoted to EFAs in the continent were under CCs (Pe'er et al., 2017), and even though the new rules in the CAP (2023-2027) may reduce the area effectively covered by CCs in some farms, incentives to shift existing farms towards more sustainable systems are likely to increase the adoption of CCs in the future (Panagos et al., 2021).

CCs are known to affect agricultural fields in multiple ways. A recent compilation of 269 studies, mainly from North America, Europe, Africa, and Asia, indicated that CCs significantly influenced 28 out of 38 physical, chemical, biological, environmental, and agronomic indicators (Jian et al., 2020). Some of the relevant benefits included the reduction of soil bulk density, water holding capacity, weed presence, crop diseases, soil erosion, runoff and nutrient leaching, and the increase of soil porosity, cation exchange capacity, soil fauna, mineralizable carbon and nitrogen, and cash crop's yield (Jian et al., 2020). Such results align with others, indicating that CCs can increase infiltration rates and attenuate soil detachment and transport (Stewart et al., 2018; Kaspar and Singer, 2015), reducing surface runoff, soil erosion, and the leaching of nutrients (Stewart et al., 2018; Olson et al., 2014). Attempts to quantify such effects at the field scale indicate that the reduction in erosion rates can vary from 15 to 23%, depending on the cropping system under consideration (Panagos et al., 2015c). Plots with CCs were also reported to have more microbial biomass and fewer weeds compared to

plots left fallow or bare (Shackelford et al., 2019), and the increase in albedo during winter months can potentially lead to the mitigation of 15.91 gCO2-eq/year/m^2 in Europe, with regional variations (Carrer et al., 2018). All the examples mentioned above illustrate how CCs can have potentially synergistic effects, as also observed by Garland et al. (2021), who found a joint increase in crop yield, soil functioning, and habitat provision for microorganisms after CCs adoption.

Concerning the impacts on the cycle of carbon and nutrients, CCs are reported to affect CO2, CH4, and N2O emissions significantly (Jian et al., 2020; Grados et al., 2022), although the effect on N2O appears non-significant in other works (Grados et al., 2022; Han et al., 2017). Some of this variability may come from the species adopted, which can also affect how the cover cropping scheme competes for nutrients with the main crop (Launay et al., 2022). In addition, CCs increase soil organic carbon (SOC) stocks significantly (Jian et al., 2020) with an average topsoil accumulation rate of 0.32 MgC/ha/year for 54 years after its implementation (Poeplau and Don, 2015). Even though the enhanced SOC may have the side effect of increasing N2O emissions, meta-analysis results indicate that the greenhouse gas emission reductions of CCs on average outweighs the emissions (Guenet et al., 2020). The increase in SOC stocks can be attributed to the larger carbon input to the soil and the reduced C loss over arable land that would otherwise remain fallow or bare, susceptible to erosion (Guenet et al., 2020). Most of the studies in the literature, however, focus on relatively short timescales. As noted by Guenet et al. (2020), the impact of CCs varies with the duration and the frequency of their implementation, and decades may be necessary to detect a significant increase in organic carbon concentrations (Poeplau and Don, 2015). Therefore, the effects must be assessed for longer periods, as many short-term experiments may overestimate the real impact of CCs on SOC increases (Guenet et al., 2020).

Another relevant effect of CCs concerns the changes in the delivery of eroded sediments and carbon to the oceans. Approximately 8.3 to 51.1 Pg of sediments are transferred to the world's oceans each year by erosion processes (Syvitski et al., 2005; Harrison, 1994), denuding the continent and affecting coastal ecosystems functioning, landforms evolution, and biogeochemical cycles (Walling, 2006). Organic carbon, for example, can be transferred to the ocean in particulate (POC) form, i.e. leaf litter, debris, and soil organic matter, or in dissolved forms, i.e. soluble particles from the decomposition of eroded organic matter (Lal, 1995). The transfer of carbon to the oceans is considered critical for the proper constraint of biogeochemical land surface models (Blair and Aller, 2012), since they affect carbon stocks

over different timescales: in the short scale, for example, the exchange of carbon between terrestrial and marine reservoirs can modulate atmospheric CO2 levels (Galy et al., 2015). The lateral transfers and the export of POC to the oceans are controlled by the SOC stocks and, mostly, by the physical processes of soil erosion and particle detachment (Galy et al., 2015), two quantities that are affected by the adoption of CCs (Stewart et al., 2018; Jian et al., 2020; Olson et al., 2014; Poeplau and Don, 2015; Guenet et al., 2020).

The adoption of CCs affects the lateral transfers of carbon in two opposing directions: the additional carbon input to the soil and the increase of SOC stocks tends to enhance the lateral export to the oceans, but the reduction in the detachment and transport of particles tends to contribute towards the opposite direction. Therefore, a relevant question is which of these two effects has a more significant influence. Answering this question, however, is not straightforward, given the lack of sediment transport and deposition movements in land carbon models. Ideally, coupled land carbon and erosion models must represent both short-term local and long-term landscape processes, with the ability to be run through a sufficiently long time range and at a sufficiently high spatial resolution to capture the effect of terrain on erosion-related processes. However, no model has achieved such goals so far due to the lack of input data, poor spatial generalization of parameters, and "immense computing power requirements" (Doetterl et al., 2016). Therefore, despite the acknowledged need to move towards more physically-based representations of erosion-related impacts on sediments and carbon transport along the landscape (Walling, 2006), large-scale land surface models often adopt simplified approaches, such as omitting the representation of transport and deposition processes (Chappell et al., 2015; Lugato et al., 2016) or setting fixed ratios for the delivery rates at the grid-cell or catchment scales (Lugato et al., 2018; Wang et al., 2017). On the other hand, process-oriented carbon erosion models compute a more realistic routing of the sediment along the landscape, but they are often limited to small domains (Nadeu et al., 2015) or to lumped simulations and short time scales when upscaled for large domains (Walling, 2006).

In this work, we focus on evaluating how CCs affect the carbon cycle over the European continent, by: i) improving the understanding of the dynamics of POC export to the oceans at a continental level; ii) quantifying whether the enhanced input or the erosion reduction due to CCs prevails on controlling the export of POC to the oceans; and iii) quantifying how a hypothetical policy of widespread CC adoption could affect the SOC budget. To do so, we first quantify the lateral transfers of carbon and sediments at the continental level from

early industrial levels (i.e., 1860) to 2050 at a daily temporal resolution with monthly forcings and 10 x 10km spatial resolution. Then, we simulate two scenarios of CCs adoption under the current climate and land use conditions with the recently developed CE-DYNAM model (Naipal et al., 2020; Fendrich et al., 2022a), which couples a spatially-explicit routing scheme to the carbon cycle of the detailed biogeochemistry land surface model ORCHIDEE (Krinner et al., 2005) to represent both ecosystem carbon fluxes and lateral movements in a spatially distributed manner (see Appendix). The strength of the approach used comes from the combination between empirically calibrated erosion rates, detailed ecosystem carbon exchange, and process-based lateral fluxes, which vary according to the terrain geomorphology and changes on the climatic and land use forcings, as well as on management activities represented through spatially-explicit maps of cover crops adoption.

4.2 . Methods

4.2.1 . Modeling

CE-DYNAM couples an emulator of the land-surface model ORCHIDEE (Krinner et al., 2005), the RUSLE-2015 (Panagos et al., 2015e) erosion model developed for Europe and adapted to include carbon erosion, and a sediment routing scheme describing the lateral movement of eroded soil and carbon in the landscape, including the transfer of particulate organic carbon to the ocean. The model was set up to run in a grid of 10 × 10 km, consistent with the highest resolution available for climate reconstruction datasets (IPSL/LSCE, 2021), which represents a compromise between the fine scale of hydrological processes and the large scale of most carbon models. A calibration was done to ensure that simulated values of sediment in rivers approximate both sediment discharge observations in river stations from the GEMSTAT database (United Nations Environment Programme, 2018a) and the aggregated ocean POC output values derived from [20]. The strength of the model is that climate, land cover, soil characteristics, and management practices directly affect all model components and their interactions. The model results were then summarized in terms of the exports of POC to the ocean (Figure 4.1 and Figure 4.2). More information about the model and its limitations are presented in the Appendix.

For the calculation of future simulations (Table 4.1), the target spatial distribution of CCs in CC_Current was defined to be the high-resolution observation-based map of Fendrich et al.

(2023) for 2016, resulting in a total area of 13 Mha. In CC_Theoretic, the expansion of CCs occurred over all croplands where erosion rates exceed 2 t/ha/year or located inside the Nitrate Vulnerable Zones (Kathage et al., 2022), reaching a much larger hypothetical CCs area of 118 Mha. Variations in the carbon budget and lateral fluxes were calculated for all scenarios and compared. The additional litter input from CCs, which cannot be modeled by ORCHIDEE, was diagnosed from a separate simulation with DayCent (Lugato et al., 2018), which provided spatially explicit gridded values (1 km resolution) at monthly time step. Such a simulation included CCs by adding an additional crop (i.e., permanent ryegrass) to the rotation when a period of at least 2 months was expected between the harvest of a cash crop and the sowing of the next cash crop. Our attempt to overcome the unavailability of climate and land use forcings until 2050 consisted of repeating the data for 2010-2017 in a loop until 2050.

Scenario	ETD enabled	CCs adoption	CCs spatial distribution
WithoutETD	✖	✖	-
WithETD	✔	✖	-
CC_Current	✔	✔	Figure S5 (left)
CC_Theoretic	✔	✔	Figure S5 (right)

Table 4.1: The variation of carbon (top) and sediment (bottom) export through time. The land-redistributed carbon can be either buried, laterally displayed or respired (for carbon). The total is the sum of redistributed and exported, and the delivery rate (DR) is calculated as the fraction of exported over the total flux redistributed on land and exported to ocean.

In order to isolate the impact of CCs, two analyses were made. The first analysis consisted of pooling together the results of CC_Current and CC_Theoretic to calculate the relative increase of their SOC stocks and the changes of ocean exports compared to simulation WithETD. We opted for pooling these two CCs scenarios for this analysis to search for common patterns that appear despite their different spatial distribution (AP B.5). Then, we calculated the model's response on SOC stocks (Figure 4.3) and on the lateral fluxes to the ocean (Figure 4.4). For the SOC stocks, calculations were grouped per classes of CCs application at the pixel level. The rate of SOC change per year was then calculated to assess the variation of the CC impact over time. For the ocean exports of POC, the results were grouped according to the average share of CC application on each basin. The second analysis consisted of calculating

the SOC budget for each scenario separately (Figure 4.5), which allowed the quantification of impacts at the continental scale.

4.3 . Results

4.3.1 . Historical and present-day simulations

Our first set of results are CE-DYNAM simulations of the impacts of erosion, transport, and deposition (ETD) on the carbon cycle from 1860 to present. Those impacts include the exposure of subsoil organic carbon due to the transfer of detached particles to downstream areas and the corresponding burial of particles at the target locations. Apart from cropland management practices, erosion rates and ocean export are affected by rainfall regime variations and land cover change (Panagos et al., 2015e), with the spatial distribution of these factors playing an important role. The height of the bars in Figure 4.1 equals the total amount of eroded material in each year. The total value is split between a fraction redistributed within the land (yellow) and another part that reaches the oceans (blue). The orange line depicts the evolution of the delivery rate delivery rate (DR, defined as the share of the eroded soil material flowing to the ocean). Our results of Figure 4.1 indicate that the DR ranges from 14.9 to 19.9% for carbon, and from 11.1 to 20.8% for sediments, respectively. These values are similar to those reported in other large-scale studies, of 10% on a global scale (Lal, 2003) or 11% and 15.3% (Borrelli et al., 2018a; Lugato et al., 2018) for Europe. For both sediment and carbon, a peak of DR between 1940 to 1960 coincides with a period where low erosion rates happened (Fendrich et al., 2022a). This pattern indicates the existence of a time delay in the response of lateral movements to a reduction in the erosion rates.

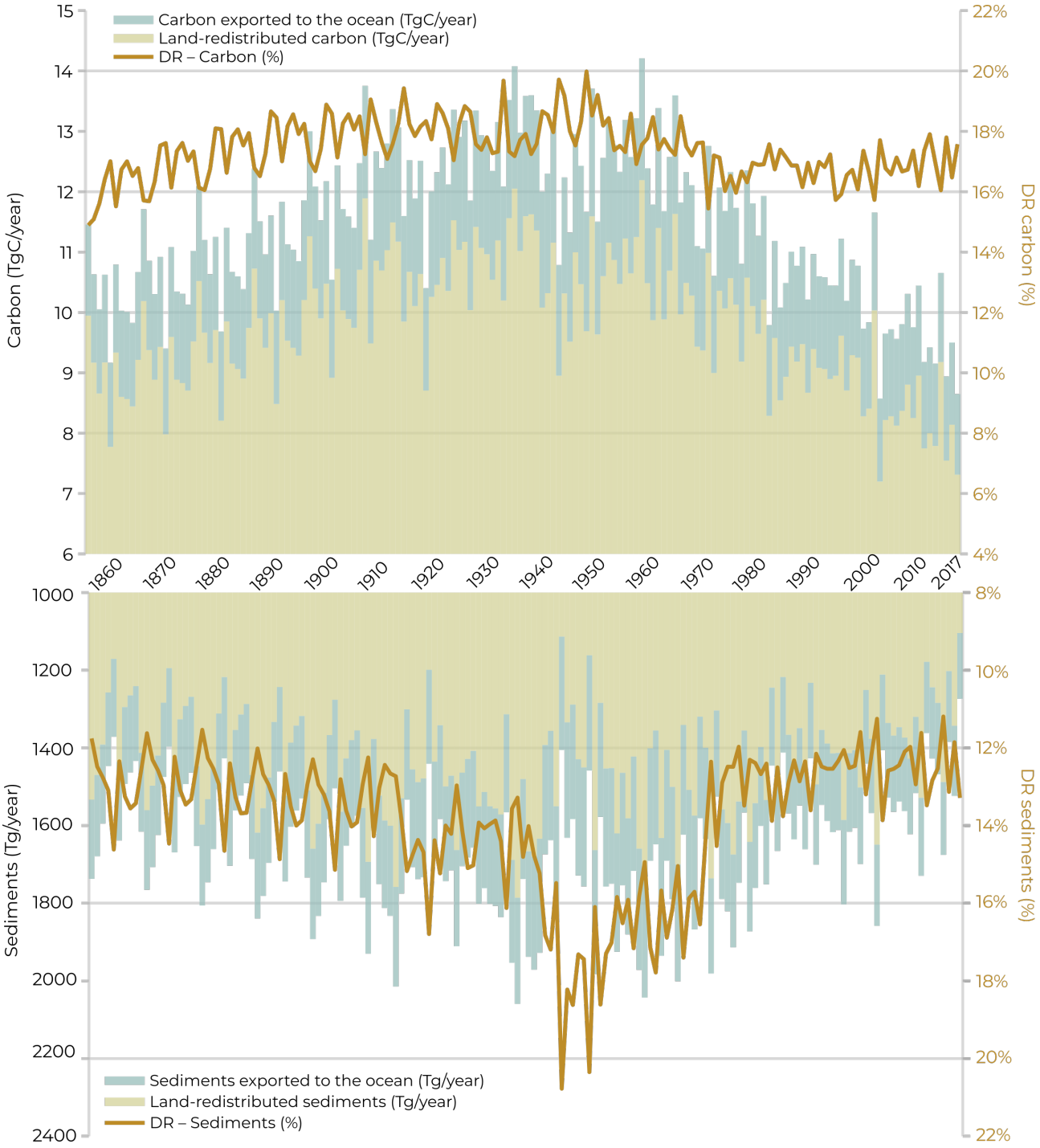


Figure 4.1: The variation of carbon (top) and sediment (bottom) export through time. The land-redistributed carbon can be either buried, laterally displayed or respired (for carbon). The total is the sum of redistributed and exported, and the delivery rate (DR) is calculated as the fraction of exported over the total flux redistributed on land and exported to ocean.

The map of lateral fluxes (Figure 4.2) shows the local imbalance between erosion removal, transport, and carbon export to the oceans. Since most CE-DYNAM cells simultaneously gain and lose carbon during the lateral transfer, areas with a higher net loss (i.e., in red tone) correspond to those where the topsoil removal by erosion exceeds the gains of sediment material from upstream sites. The figure also shows the magnitude of the flux of POC export to the oceans, represented with grey circles. It can be seen that most carbon lost to the ocean comes from a reduced number of regions, namely Great Britain, Italy, Greece, the Balkan States at the Adriatic Sea, and the south of Spain. In all cases, the regions belong to the Mediterranean or North Atlantic basins (see Appendix), inducing their large ocean export. Out of the total average of 1.95 TgC/year exported for 2000-2017, the contribution of each group of basins corresponded to 46.09% for the North Atlantic Ocean, 43.44% to the Mediterranean, 8.06% to the Baltic Sea, 1.98% to the Black Sea and 0.43% to other regions (AP B.1). An underlying process for those high carbon exports is the proximity between regions with high erosion rates and the coast. This result suggests that catchment elongation plays an important role in controlling ocean export. Another element that supports this explanation is that the opposite effect can be perceived in the Black Sea, where the high losses in Austria, Romania, and Bulgaria do not necessarily convert into a high export to the sea for these basins. However, other factors such as landscape connectivity can not be discarded since their effect is not properly captured at the spatial resolution adopted, i.e., 10 x 10 km.

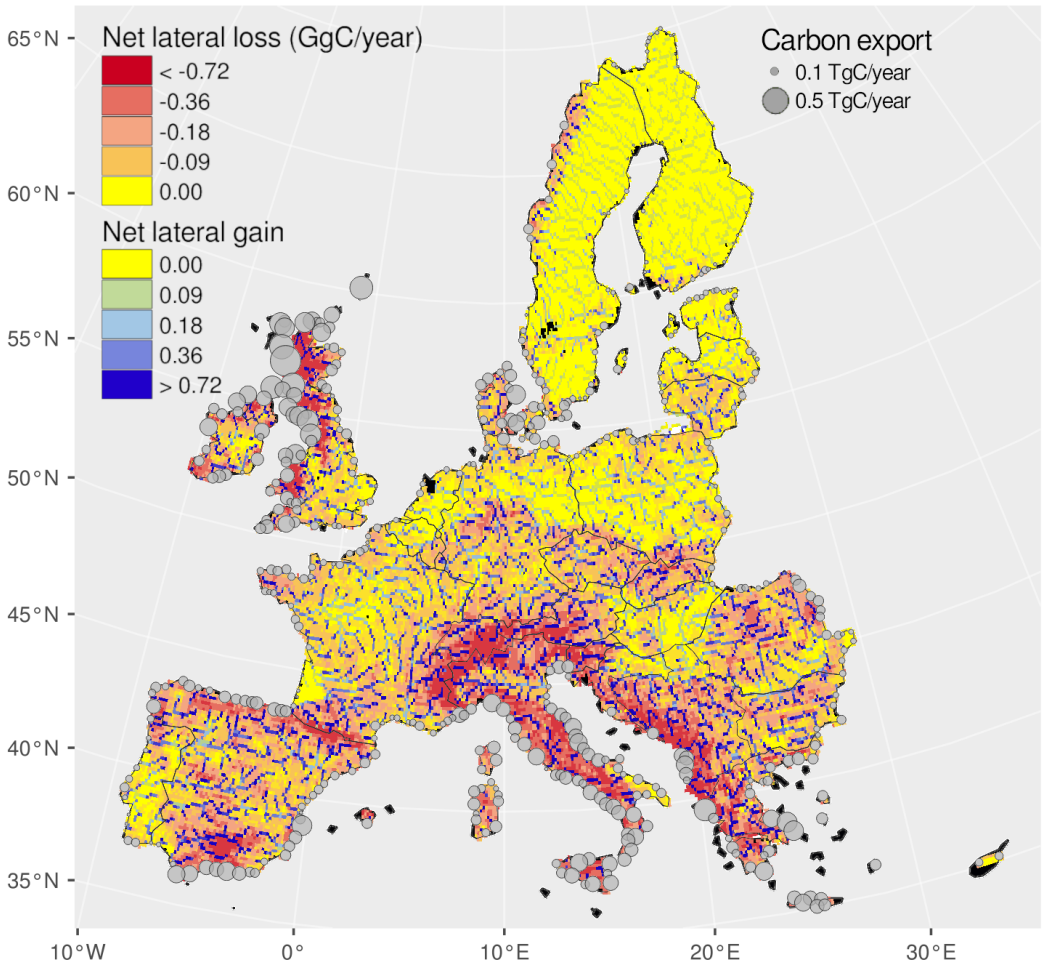


Figure 4.2: Net lateral carbon transport and carbon export to the oceans, average for the period 2010-17. Results refer to the simulation WithCC with enrichment factor equal to 1 (see Appendix). In order to enhance visualization, the diameter of each circle is proportional to the carbon export raised to the power of 1/3. Carbon export ranges from 0 (no circle) to 0.5 (largest circle) TgC/year.

4.3.2 . The future impacts of cover crops

Four scenarios were designed to evaluate the impact of ETD with and without CCs on soil carbon fluxes and stocks until 2050. In scenario WithoutETD, a default land surface model simulation was used, therefore not including ETD or CCs. Then, WithETD included the soil erosion rates and the ETD processes but no CCs. The comparison of the results of With-

outETD and WithETD allowed us to isolate the effect of ETD processes. Then, the scenario CC_Current consisted of WithETD plus a realistic application of 13 Mha of CCs. Finally, scenario CC_Theoretic assumes a maximum CC expansion up to 118 Mha based on Panagos et al. (2021), corresponding to a strong policy incentive to reduce soil losses of soil and prevent eutrophication (Shackelford et al., 2019; Jian et al., 2020; Martin, 2019). Table 4.1 in the Materials and Methods section summarizes the characteristics of each scenario, and more information can be found in the Appendix.

Figure 4.3 shows the relative change of SOC stocks during the CCs expansion period. The left image shows the cumulative increase, while the right one shows the annual SOC rate of change for different classes of CC application in percent coverage of arable land. The continuous lines representing the averages indicate higher increases in SOC stocks for higher rates of CC application. The uncertainty bands presented correspond to the standard deviation of 15 sensitivity simulations run (see Appendix). The increase of SOC induced by CCs is more uncertain for CC fractions below 30%. Above this level, the number of pixels in CC_Current is low, leading to a lower inter-scenario variability and therefore to narrower uncertainty bands. For CC fractions higher than 30%, the average additional increase of SOC in the CC scenarios compared to WithETD reaches 34.5 ± 0.4 ‰ after 50 years. Such an increase is, however, not linear. The yearly average rates of change show that the SOC sequestration rate induced by CCs increases to reach a maximum after 10 years, and then declines. This means that SOC continues to accumulate, albeit at a slower rate, even after CC has reached its maximum expansion.

63.5 ± 2.4% out of the total additional storage in Figure 4.3 happened after 2017, when the fraction of CCs had already reached its maximum expansion. For pixels with high CC fractions, a peak of additional SOC sequestration rate at 1.3 ± 0.1‰/year was obtained around 10 years after the beginning of their application. After the peak, the sequestration rate decreases gradually until the end of the simulation period. Combined, the narrow uncertainty band for such a result and the fact that the CCs scenarios are pooled together indicate that the effect is significant and does not depend on other factors that vary spatially across the continent, such as climatic and biological elements that affect the SOC cycle. Overall, applying CCs at the current rate would have a maximum increase of 0.03‰/year and applying them at a maximum rate would peak at 0.19‰/year. In a country level, the two countries with the overall highest increases in CC_Current would be Denmark and Poland, with peaks of 0.23 and 0.20‰/year, respectively, while in CC_Theoretic those countries would be Hungary and

Italy, with peaks of 1.23 and 0.84‰/year, respectively.

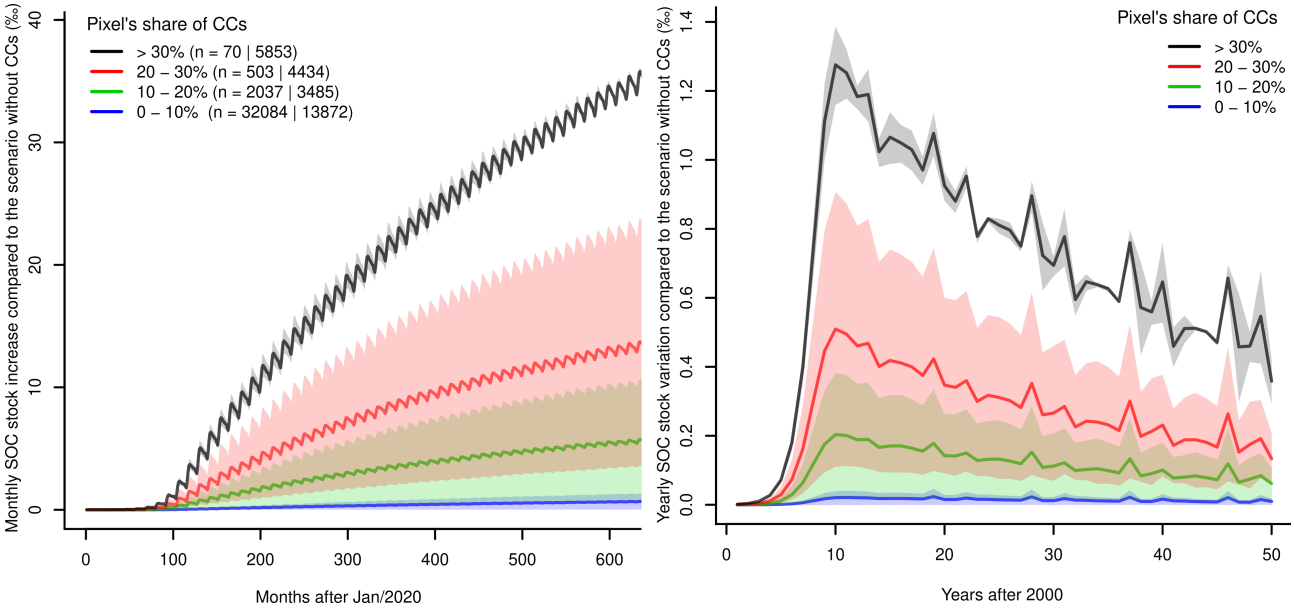


Figure 4.3: SOC stocks variation as a function of the pixel's share of cover crops (CCs): relative cumulative increase per month (left) and annual changes (right). The bands around the mean are the standard deviation of the 15 parameters' sensitivity simulations (see Appendix). The number of pixels in each class is reported as: (n = <number in CC_Current> | <number in CC_Theoretic>).

For ocean export, Figure 4.4 shows a clear inverse relationship between the average rate of CCs application and the carbon export to the oceans. Analogously to the case of SOC stocks, the decrease in carbon export stabilizes at some point. After around 13 years of CCs application, the decrease of C export stabilizes at around -43.1 ± 24.8 ‰ compared to the WithETD scenario when the basins contain more than 15% of CCs in croplands. It must be noted that, similarly to the SOC stocks, the classes below 15% have narrower uncertainty bands due to a lower inter-scenario variability. The results indicate a delayed response of the ocean export compared to the 10 years response time observed for the peak of SOC storage. When the scenarios are considered separately, the same plateau is reached on the class above 15% of CCs in croplands, but amounting to -20.9 ± 4.1 ‰ in CC_Current and -65.3 ± 12.5 ‰ in CC_Theoretic.

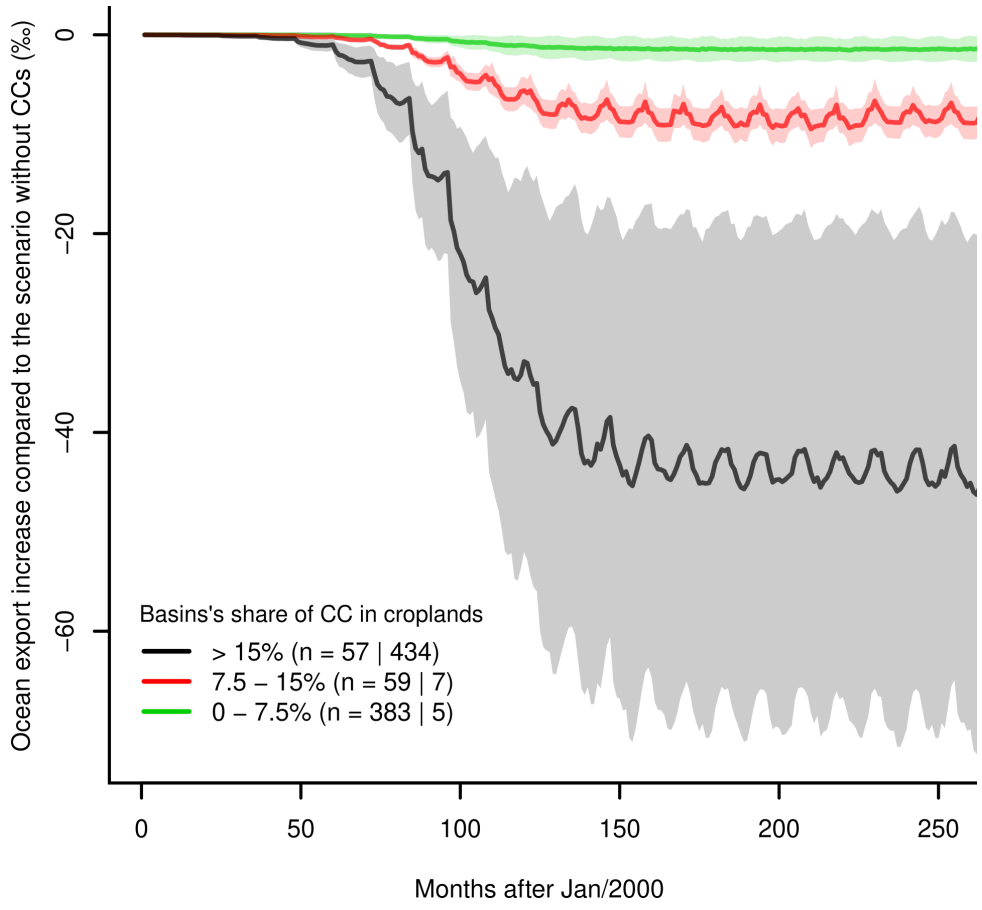


Figure 4.4: Carbon export variation as a function of basins' share of cover crops (CCs). The bands around the mean are the standard deviation of the 15 simulations incorporating uncertainty on model parameters (see Appendix). The period corresponds to 2000 to 2020, the period 2021-2050 was removed due to lack of relevant changes in the pattern displayed. The number of basins in each class is reported as: (n = <number in CC_Current> | <number in CC_Theoretic>).

The SOC budgets for all scenarios in Figure 4.5 show that, compared to WithoutETD, WithETD has a slightly lower soil respiration rate (933.22 vs. 932.75 TgC/year) due to the continuous removal of carbon from the topsoil by ETD processes. Scenario WithETD also sets up reference values that can be compared against CC_Current and CC_Theoretic. The gross eroded carbon in WithETD, CC_Current and CC_Theoretic are 14.13, 14.06 (-0.5%) and 13.09 TgC/year (-7.4%), indicating a more considerable decrease when more CCs are adopted. The same relationship is found for ocean export, which remains equal to 1.95 TgC/year for WithETD and CC_Current but decreases up to 1.84 TgC/year (-5.6%) in CC_Theoretic, and for burial and sub-

soil exposure. The former varies from 12.18 TgC/year in WithETD to 12.11 and 11.23 TgC/year (-0.6% and -7.8%, respectively) in CC_Current and CC_Theoretic, respectively, while the later reduces from 7.68 TgC/year in WithETD to 7.65 TgC/year (-0.2%) in CC_Current and to 7.19 TgC/year (-6.4%) in CC_Theoretic. Two fluxes that increase in CC_Current and CC_Theoretic are the litter input and respiration rates due to the enhanced input caused by adding a new species to the crop succession. In WithoutETD and WithETD, the litter input equals 989.61 TgC/year (2.01 MgC/ha/year), and increases +1.49% to 1002.37 TgC/year (2.04 MgC/ha/year) in CC_Current and +7.96% to 1065.77 TgC/year (2.17 MgC/ha/year) in CC_Theoretic.

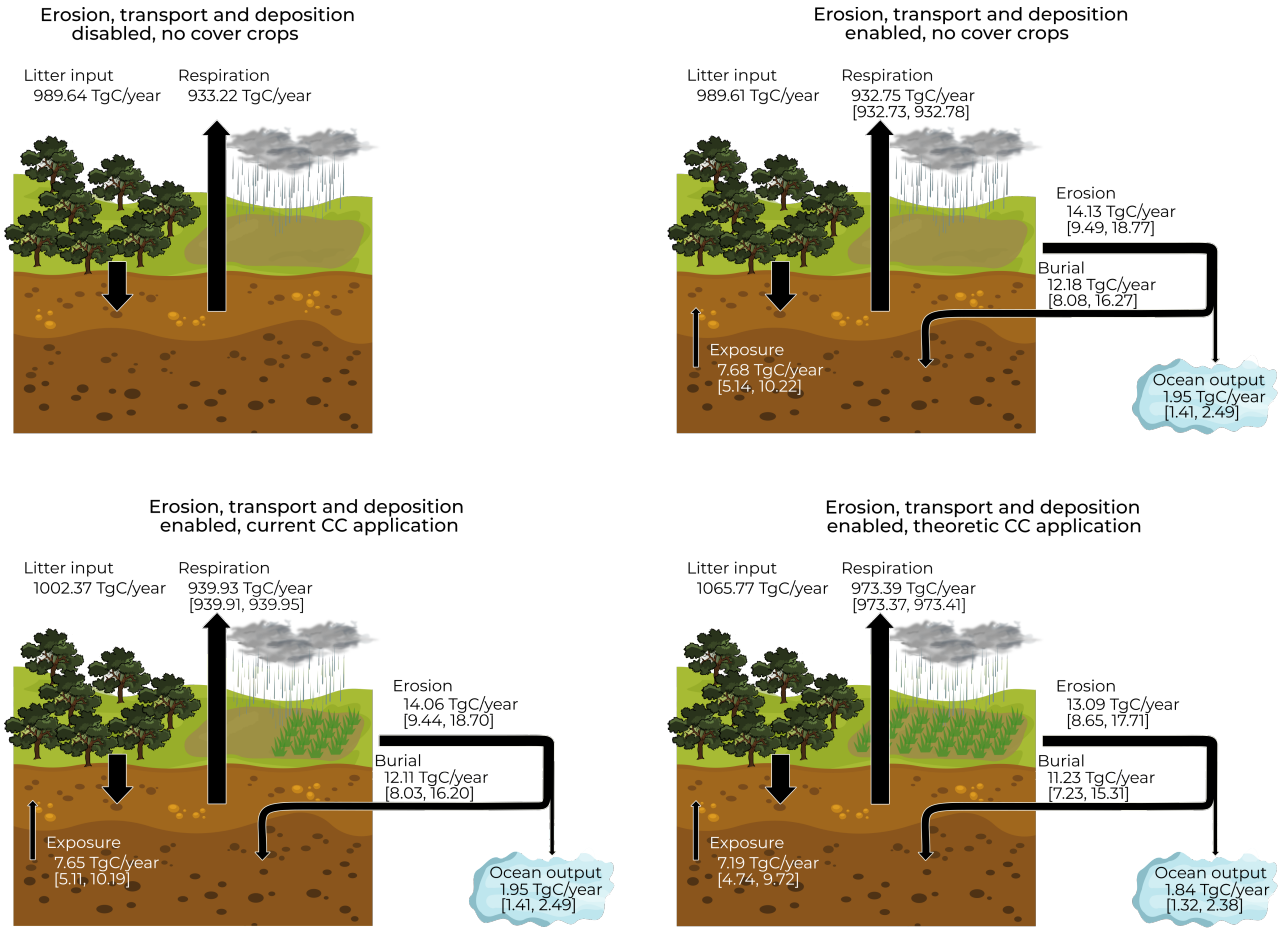


Figure 4.5: SOC budget for scenarios WithoutETD (top left), WithETD (top right), CC_Current (bottom left) and CC_Theoretic (bottom right). Values are reported as: “average [min, max]” of the 15 parameters’ sensitivity simulations (see Appendix).

4.4 . Discussion

In the historical and present simulations, the difference in range between sediment and carbon DRs (Figure 4.1) arrives from a combination of two factors. The first factor is the mismatch between areas with high erosion rates and those with high exposed SOC stocks. Figures S6 and S7 show that carbon-rich soils often have low erosion rates and vice-versa. Two clear examples are Italy, with low SOC stock and high erosion rates, and Sweden, with the opposite combination, i.e., high SOC stock and low erosion rates. Three particularly relevant regions where the agricultural area increased from 2000 to 2017, and where both SOC stocks and erosion rates are predominantly high, are Bosnia and Herzegovina, Montenegro and Albania (AP B.7), contributing to the high export to the Mediterranean Sea (AP B.1). The same pattern is also found in the United Kingdom and Ireland (AP B.7). The second factor affecting the DR of carbon distinctly from sediment is the fact that after particle detachment, the carbon transported and buried off-site is partially released to the atmosphere through respiration by microorganisms. Such an effect is particularly relevant in inland mountain regions such as Switzerland (AP B.6), which are far from the ocean and have high losses of carbon (Figure 4.2). In these regions, burial and respiration happen during the journey of eroded carbon along the landscape, before reaching the rivers and the ocean.

When cover crops are considered, the period of 13 years for POC export stabilization (Figure 4.4) indicates the presence of a delayed effect compared to the peak in SOC increase rates after 10 years (Figure 4.3). A possible explanation for such a lag can be the time needed for particles to reach the ocean after erosion events. The residence time of carbon particles varies across the landscape, as their stability relates to geomorphological characteristics such as the dominant grain size (37). The results indicate that in the tradeoff between the enhanced carbon input, which increases POC availability, versus the reduction of soil losses from erosion, which decreases POC availability, the second factor prevails over the first in controlling ocean POC exports. Considering that CC_Current and CC_Theoretic are pooled together in the analysis, such an effect obtained is persistent regardless of the spatial configuration of CCs. Other factors, such as basin elongation and land use, seem to affect only its magnitude but not its direction. Even though the model is sensitive to climatic, pedologic, and watersheds' morphological properties, all effects found seemed persistent despite the different spatial distribution of CCs application in the scenarios considered.

The increases in carbon input in the CCs scenarios (Figure 4.5) can be compared to the

target of the 4‰ initiative, which aims at increasing SOC stocks by 0.4% per year. Generally, the effects of CCs are very low compared to the estimated increase of 30-93% (38, 39, 40) necessary to achieve this target. As shown in Figure 4.3, the maximum annual contribution when the CC fraction exceeds 30% is only approximately one-third of the 4‰ initiative's target. Overall, the maximum annual increase would correspond to 0.75% of the target in the scenario with a realistic CC application and 4.75% in the scenario with a theoretic CC application. With the highest increase at the country level, Denmark and Poland would have a maximum additional SOC storage of 5.75 and 5% of the 4‰ target in CC_Current, while Hungary and Italy would reach a maximum of 30.75% and 21% of it in CC_Theoretic. Furthermore, the highest increase in inputs obtained here, of 0.16 MgC/ha/year in CC_Theoretic (Figure 4.5), is also very low compared to reference absolute values for the 4‰ initiative, amounting to only half the lower uncertainty bound estimated by Riggers et al. (2021). These results reinforce the idea that CCs are insufficient to achieve the 4‰ targets without other additional measures, such as the adoption of reduced tillage and the conversion of marginal croplands to grasslands (Minasny et al., 2017).

It must be noted that our results have known limitations (see Appendix). For example, they were generated using recent climate forcing repeated in a loop for the future, meaning that the uncertainties and impacts of future climate change are not taken into account. They also neglect that over such a long period of 50 years, cropping systems may change dramatically with effects of major importance compared to those of CCs. The results also do not quantify important co-benefits of cover crops that are likely to occur, such as the avoided losses of eroded phosphorus, the reduced nitrogen leaching and the potential benefits for crop yields. Despite these drawbacks, we attempted to capture the impact of CCs on the transport of soil particles and the SOC cycle changes in space and time, a dynamic that is not yet fully captured by most land surface models. To the best of our knowledge, our approach constitutes the first successful attempt to represent such dynamics with a spatially distributed approach at a continental level and a sufficiently high spatial resolution, allowing a more detailed understanding of the fluxes from the detachment of particles until their export to the oceans.

Therefore, despite the limitations of our work, we hope our results shed some light on the fluxes involved in carbon transport from the land to the ocean. We highlight that increased rainfall variability due to climate change may significantly impact the results reported in this work, potentially creating lateral carbon fluxes in the same order of carbon farming activities.

Increasing the knowledge on the integrated carbon cycle is fundamental to properly separate real net removals from the confounding effect of carbon displacement.

4.5 . Learnings

This chapter evaluated the impacts of different cover crops scenarios on the SOC stocks and POC export to oceans in Europe. The results led us following learnings:

- The delivery rate of sediments presents a different variability and range than carbon due to spatial mismatch between erosion rates and carbon stocks.
- Cover crops tend to increase SOC stocks even after their application reaches a maximum. After this moment, however, the accumulation rate decreases.
- The peak in SOC accumulation happens before the stabilization of ocean export, indicating a delay in the former process. The delay can be explained by the time taken by particles to travel across the landscape after erosion events.
- The magnitude of effects is low if compared to the targets of relevant policy initiatives such as the 4‰ target. The highest increase simulated reached only half the lower uncertainty bound found in the literature for achieving the target.
- The spatially explicit representation of carbon movement across the landscape attempts to increase the realism of land surface models by reducing existing confounding effects.

Before this chapter, this thesis had two parallel development lines: the mechanistic approach for CE-DYNAM development and the empirical modeling for cover crops mapping. This chapter attempted to combine these lines to produce meaningful simulation results to understand better the impacts of cover crops and soil erosion in the carbon cycle.

5 - Conclusions and perspectives

5.1 . Conclusions

In this thesis, we upscaled and calibrated CE-DYNAM over the European continent to allow its use as a tool to inform agricultural decision-making. In the first chapter, we re-designed the CE-DYNAM model with a focus on performance and scalability, including a calibration procedure to make its predictions more realistic. The new version (v2) greatly improved runtime and memory usage, making continental-level computations feasible. According to our literature review, the computational problem we addressed is one of the major challenges of coupling erosion to the soil organic carbon cycle. In the second chapter, we focused on land management, the interface between erosion and carbon. Since a common problem of representing land management, specifically cover crops, in models is the need for spatially explicit data about their occurrence, we attempted to disclose the fine-level patterns hidden in the aggregated-level statistics provided to the public by the European statistical agency. The results constitute the first-ever map of cover crops produced at the European level, validated for France using parcel-level information. Finally, in the third chapter, we improved the physical realism of CE-DYNAM to its third version (v3), and used it to reconstruct the historical export of particulate organic carbon to the oceans and to evaluate the impacts of two cover crop scenarios on the European carbon budget.

These works led us to the following conclusions:

- Given the target spatial resolution of this work, 10km, the matrix representation of lateral fluxes could upscale CE-DYNAM up to the continental level. However, the dimension of the resulting sparse matrix equals the product between the number of grid cells, plant functional types, vertical soil layers, and soil carbon pools. This means that going from the continental to the global scale would directly increase complexity, and the application would only be viable by simplifying the other model components or with further model developments. We opted for coupling CE-DYNAM to ORCHIDEE, a model that has the particular characteristic of representing land cover as fractional, with n proportions summing to 100%. In DayCent, land cover is represented with a sin-

gle map, which means that connecting CE-DYNAM to DayCent is a potential way to run the model in a $\sqrt{1/n}$ times higher spatial resolution without increasing complexity.

- The results of the first chapter allowed us to understand the dynamics of soil erosion since the early-industrial age. Coupling historical climate and land cover forcings to the erosion model yielded present-day erosion rates with a positive correlation against other modeling studies and local- and national-scale predictions. When extrapolated to the past, the same model described an increase of the aggregated rates from 3.18 t ha^{-1} year^{-1} in the decade 1860-1870 to a peak of 3.76 t ha^{-1} year^{-1}, followed by a decrease to 3.04 t ha^{-1} year^{-1} in 2010-2018. The individual analysis of the driving factors allowed us to identify the land cover as a driver of the long-term trends in the erosion rates and the rainfall variability as a driver of the short-term fluctuations.

- The results of the second chapter allowed us to verify that the satellite data from Sentinel 1, although not optical, can be useful in identifying the existence of cover crops on the field due to its high temporal and spatial resolution and the lack of gaps on the time-series. The disaggregation model could properly identify the difference between cover crop and winter cash crops in several situations in France, which indicates that it was able to capture the proper phenology patterns. The spatial pattern of cover crops at 100m in Europe allowed us to find where cover crops are more widely adopted, namely the North of France, Germany, and Denmark. The production of this dataset constituted a new piece of information to be used by future modeling studies, and the validation for other countries remains subject to further work.

- In the third chapter, we generated the sediments and carbon delivery rate from the European countries to the oceans since the early-industrial age. The data suggested a higher variability in sediment than carbon export, indicating that other processes, such as burial and respiration by microorganisms, have a significant impact on the eroded carbon before its export. The results also showed that the increase in soil organic stocks induced by cover crops tends to reach a maximum after 10 years of their adoption, and then decrease. The export of particulate organic carbon to the ocean decreases in all cases, reaching a minimum after 13 years of the adoption of cover crops. The difference between these two values (i.e. 13 and 10 years) indicates that the ocean export has a lagged effect due to the time needed for sediments to cross the landscape and reach the ocean. At the continental scale, we found that most of the carbon exported comes from a limited number of regions, including the United Kingdom, Italy, Greece, and the

Balkan States.

- The results of the third chapter also indicated that the modifications on the carbon budget induced by cover crops are relatively small and not sufficient to achieve the targets of the 4‰ initiative according to the current literature. Overall, the current rate of cover crops would contribute to a maximum of 0.75% of the target, and a hypothetical scenario of full application would peak at 4.75% of it. Together, these results indicate that additional management measures are necessary to reach the 4‰ goal. These results, however, do not include the impacts of climate change due to the unavailability of forcings in the relatively high resolution used.

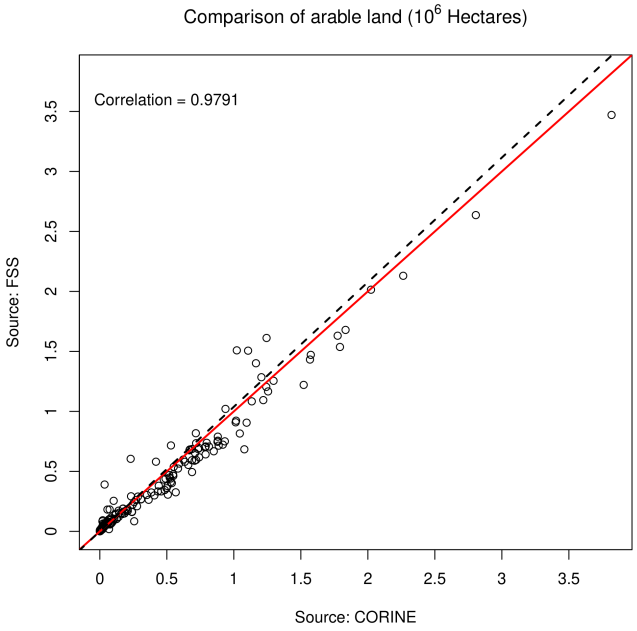
5.2 . Perspectives

Despite the progress made during this thesis, further work is necessary to address the following issues:

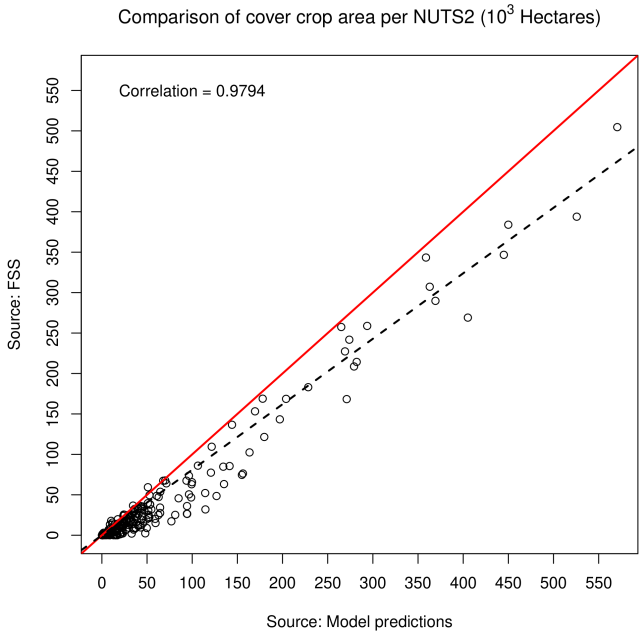
- Even in its final version, CE-DYNAM still does not represent some important carbon processes related to soil erosion, namely the greater susceptibility of the soil organic carbon's light fraction to the detachment of particles, the breakdown of particles during transport and the corresponding reaggregation during deposition. The final version also does not include the conversion from particulate to dissolved organic carbon during transport, nor the specific dynamics of dissolved organic carbon. Dams are another missing component of the river network that can have a significant impact on the model's results.
- The map of cover crops generated is only validated using parcel data in France due to the lack of publicly available data in the other European countries. More information is expected to be released to the public in the future, which could lead to better disaggregation results. Therefore, validating and updating the current map could be done as further work.
- Other aspects of land management must be better developed in CE-DYNAM, including the gradients of residues management and tillage adoption. These topics present similar challenges as cover crops in terms of data availability, and have received growing attention in environmental policies.

- Both erosion and cover crops affect the water holding capacity of a soil, which in turn impact its productivity. Currently, a known limitation of CE-DYNAM is that it lacks a proper representation of the impacts on the net primary productivity (NPP), which can greatly affect the results. During the development of this thesis, it was understood that investigating the relationship between erosion and NPP would be a work on its own, and that can be addressed by further work.

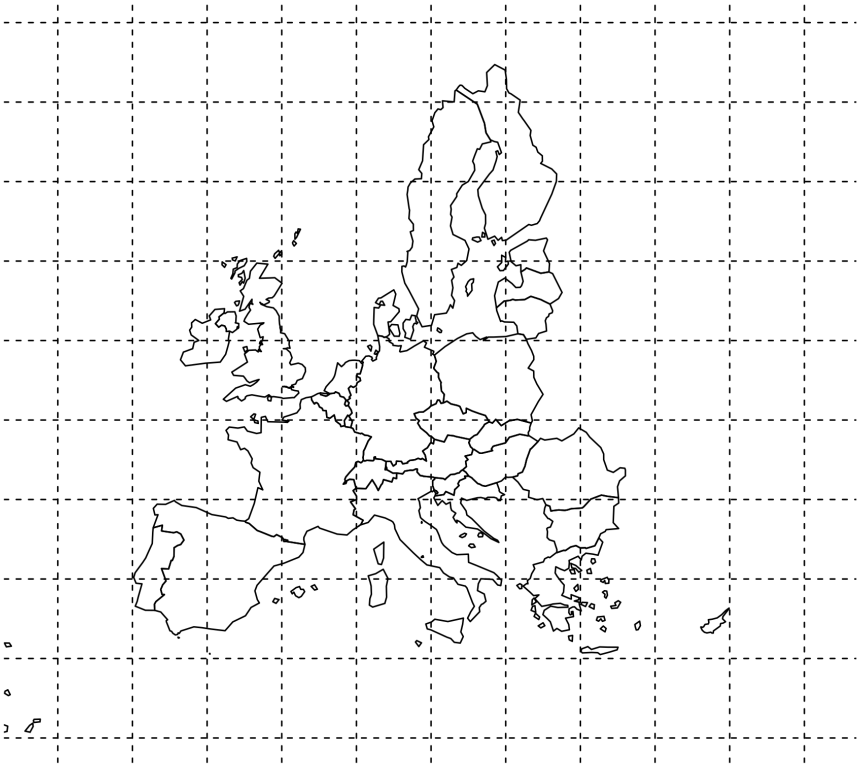
A - Appendix to: From regional to parcel scale: a high-resolution map of cover crops across Europe combining satellite data with statistical surveys



AP A.1: Comparison between the amount of arable land in Corine and the Farm Structure Survey (FSS)



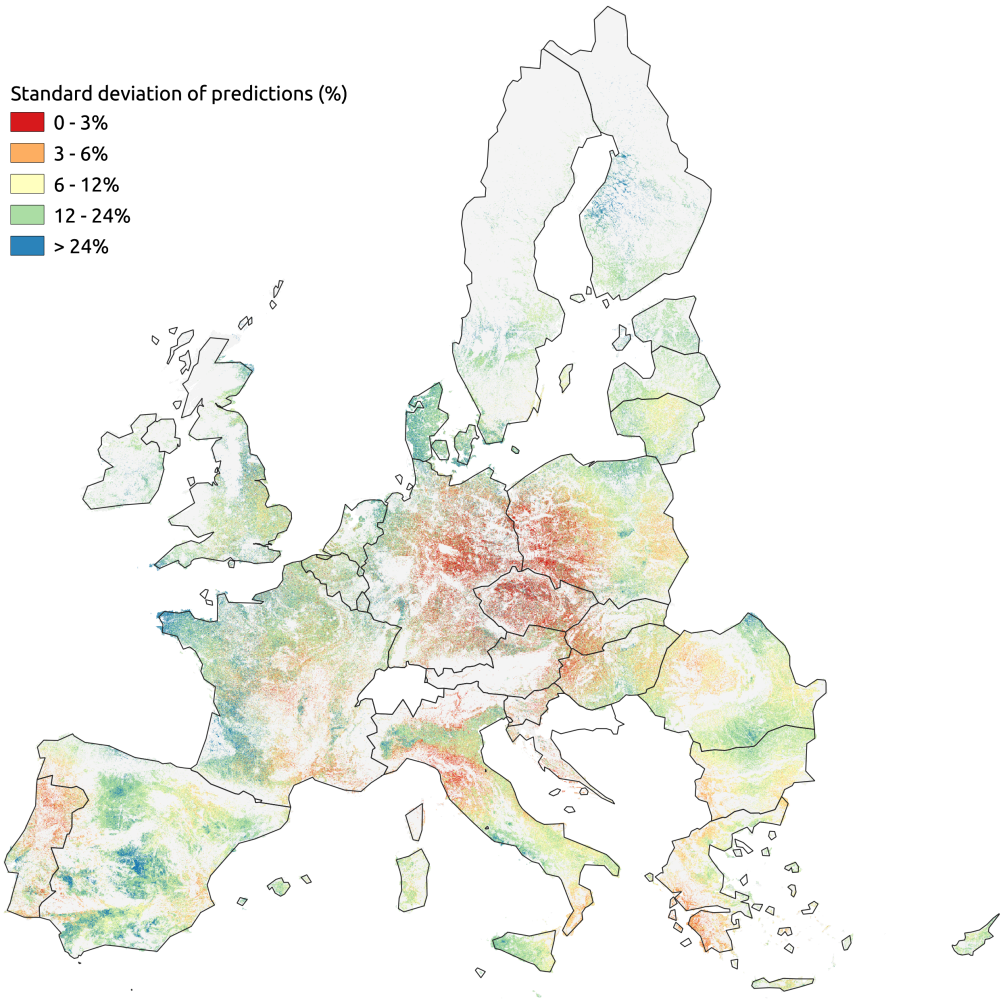
AP A.2: Comparison of the total area of cover crops predicted by the model and reported by the Farm Structure Survey (FSS)



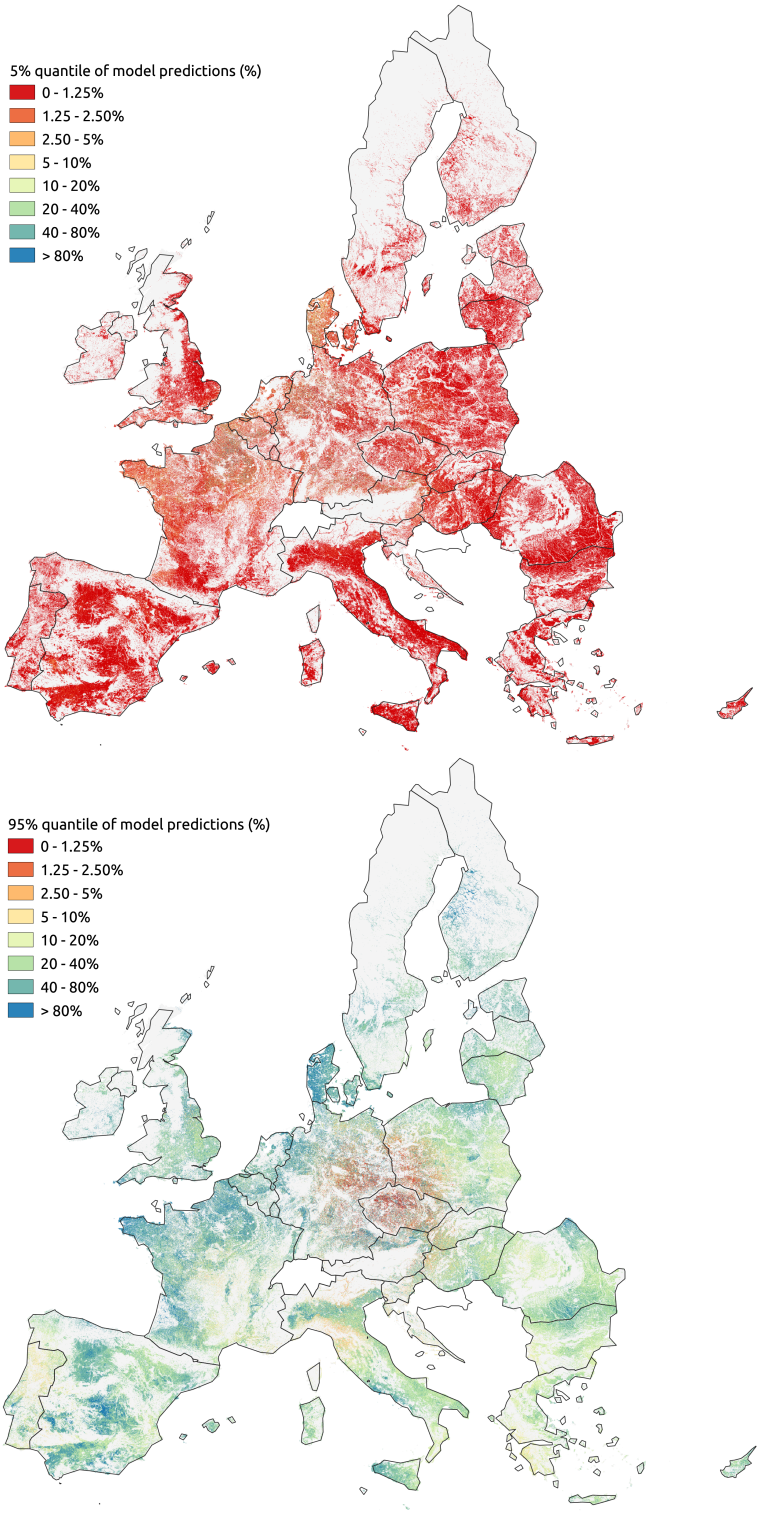
AP A.3: Knot placement for the spatial basis used in the current work

	Option #1	Option #2	Option #3	Option #4
Overall	0.730	0.720	0.741	0.727
West	0.713	0.707	0.748	0.703
North	0.734	0.752	0.769	0.755
East	0.741	0.679	0.696	0.727
Center	0.741	0.734	0.748	0.747
South	0.659	0.669	0.654	0.630

AP A.4: Four best models found during the model selection procedure, and their regional Area Under the Curve (AUC) values. Option #3 was chosen and presented in this work.



AP A.5: Model uncertainty: standard deviation of model predictions. Values calculated using the smoothers' estimated variance-covariance matrix.



AP A.6: Model uncertainty: 0.05 and 0.95 quantiles of model predictions (top and bottom, respectively). Values calculated using the smoothers' estimated variance-covariance matrix.

B - Appendix to: Improving land management representation in a constrained European carbon scheme with lateral displacement to oceans

Model framework

In this work, we use the CE-DYNAM (v3) model (Naipal et al., 2020; McClelland et al., 2021) to represent the feedback between the carbon cycle and erosion, transport, and deposition (ETD). The model contains three modules that interact: a carbon cycle emulator, an erosion model, and a sediment routing scheme. A formal description of model v2 equations can be found in Fendrich et al. (2022a). The description of v3 below, its parameters, and its differences against v2 is written in plain language without the long mathematical formulation.

The carbon cycle emulator used was an offline version of the outputs produced by an ORCHIDEE 2.2 (Krinner et al., 2005) run at a 0.125-degree spatial resolution (i.e., approximately 10km at the Equator) from 1860 to 2017. When CE-DYNAM runs without the two ETD modules, its results coincide with those produced by ORCHIDEE. When ETD is enabled, the carbon is affected in multiple ways. First, the land use and rainfall precipitation lead to erosion rates calculated using RUSLE2015 (Panagos et al., 2015e). Such rates induce two in-site fluxes: one for topsoil carbon removal and another for subsoil exposure. Next, off-site effects are also induced. Depending on the terrain, the detached carbon sediment is routed downstream across the landscape, and when an area receives an incoming flux of carbon from an upstream region, the carbon can be either buried or re-routed. After several timesteps, the carbon, once buried, can be re-exposed. The terrain is preprocessed to fill routing sinks, i.e. pixels in the middle of the continent that block the sediment from flowing downstream, in order to ensure hydrologically correctness.

In comparison to its previous version (v2), CE-DYNAM v3 presents several modifications, which include the reduction of redundant and duplicated fluxes, the enforcement of mass conservation, and the removal of the distinction between hillslopes and valley bottoms using terrain data, which proved unhelpful at the desired spatial resolution. Three major modifications were also made to the model structure and calibration. The first consisted of removing the vertical soil discretization from the calibration procedure. The second consisted of removing splines to reduce particle detachment, an artifice used to facilitate calibration in CE-DYNAM v2. In CE-DYNAM v3, particle detachment is only controlled by the erosion rates, and what is subject to calibration is the intensity of the lateral movements. The third change consisted of modifying the calibration procedure by changing the model's objective

function. In CE-DYNAM v2, the observations of sediment flux were compared to CE-DYNAM's stocks, which led to overly optimistic and inflated results. In CE-DYNAM v3, the model's lateral sediment output is set to match the observed data. With these changes, more physically-consistent results and non-inflated metrics could be produced.

Model parameters and calibration

CE-DYNAM v3 contains three components for calibration. The first component is a curve relating the flow accumulation to the share of sediments that are routed downstream, which we assume to be represented with a generalized logistic curve. The second and third components consist of the intensity of the lateral and vertical fluxes, which can be easily controlled thanks to the matrix representation of fluxes developed by Fendrich et al. (2022a). In order to find the best set of parameters for the three components, the calibration procedure consists of running the routing scheme for sediments only and comparing the lateral fluxes predicted against sediment discharged measured data.

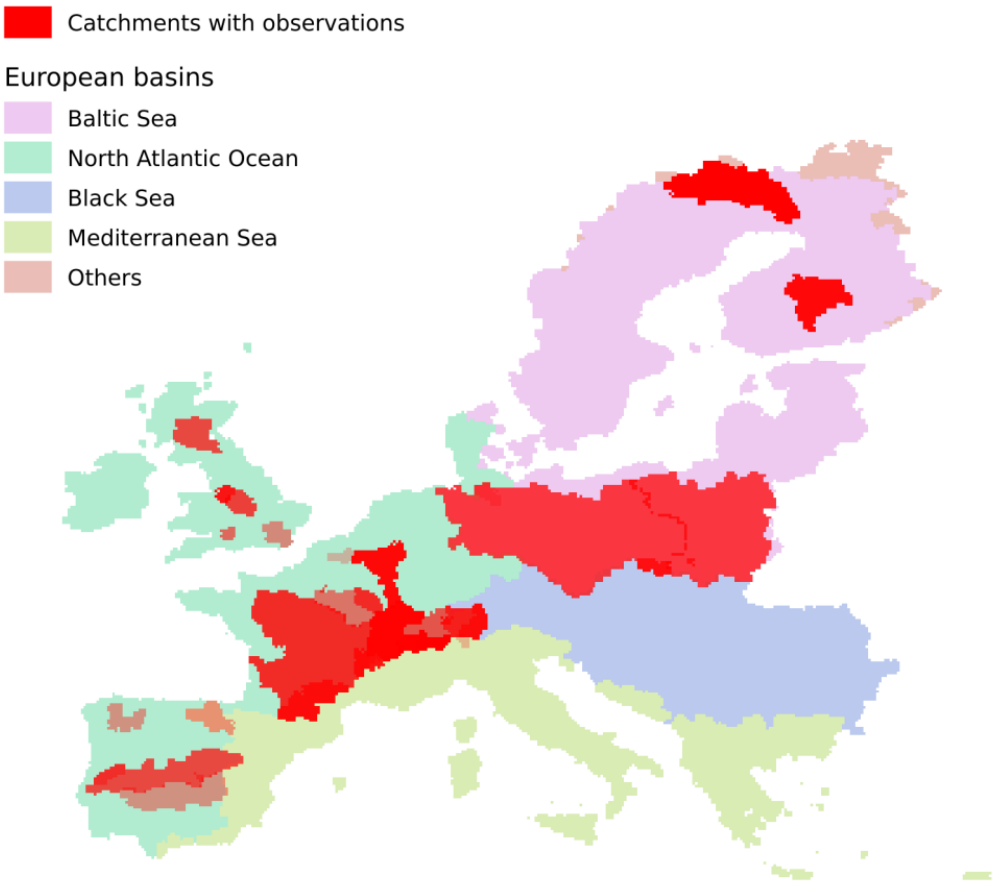
However, since several combinations of parameters can lead to acceptable results, an additional constraint is added: that the amount of sediment exported to the ocean by the basins approximate as much as possible those values predicted by Borrelli et al. (2018a). This additional constraint adds complexity by converting the problem from a single- to a multi-optimization problem. Two further complications of model calibration are the domain of the variables and the computational complexity. Even though the generalized logistic curve only contains real-value parameters, the intensity of lateral and vertical fluxes are controlled by integers, which precludes the use of many methods that require continuous and differentiable functions. Besides, one single function evaluation for sediments only takes several hours in an Intel(R) Xeon(R) CPU E5-2640 v4 2.40 GHz processor, which complicates the process of finding a solution.

Given all the complications mentioned above, we opted to simplify the function evaluation and optimize it using a random search across the domain of parameters. Since the ultimate goal of calibration is to apply the lateral movements for carbon simulations, the calibration with sediments replicates the same routing scheme behavior. Therefore, a complete function evaluation consists of: i) generating monthly model matrices (see Fendrich et al. (2022a)) as a function of the set of parameters; ii) calculating equilibrium for the early-industrial period; iii) running the model monthly from 1860 to the date of observations; iv) calculating difference between predictions and observations. The simplified procedure cuts simulation time by modifying the temporal resolution. This definition leads to the following procedure: i) generating annual model matrices; ii) calculating equilibrium for the early-industrial period; iii) calculating the difference between predictions at equilibrium and observations. Thus, due to the computational complexity of the process, the simplified calibration can be seen more as an attempt to lead the model execution toward the correct magnitudes

than to predict the exact observed value in the exact observed date. The simplifications reduced the running time from several hours (complete version) to two hours (simplified version).

For the random search, we randomly sampled and evaluated 1350 points in parallel, and for each simplified evaluation, we reported the absolute error between model predictions vs. observations for both: i) sediment discharge measured data and ii) aggregated ocean exports. Then, for the multi-optimization solution, we used an interactive a posteriori method (Miettinen, 1998), which relied on manually inspecting the group of possible solutions (i.e., that of low absolute error for both objective functions) to choose the most physically-relevant set of parameters. The choice of an interactive over a non-interactive method proved necessary due to the small number and variable range of the aggregated ocean output data (ii), which created a problem: in several cases, a low absolute error was obtained when the model predicted correctly the observation with the largest magnitude, but zero for all the other basins. Even though the use of the absolute error improved this issue compared to other alternatives, for instance the squared error, it could not solve it completely.

The sediment discharge observations used consisted of observations of total suspended solids and river discharge made available by GEMStat for Europe. The preprocessing step consisted of first filtering only the catchments whose reported catchment area is similar to that calculated using our digital elevation model and then calculating yearly averages. The dataset used consists of 372 yearly averages on 38 European rivers from 1979 to 2003. From this total, 31 were used for model fitting and 7 (around 18%), for validation. Even though the simplified calibration procedure can not capture multi-year variation, we retained the repeated measures in the dataset to give more weight to the basins where more observations are available. The 38 catchments cover 27.18% of the study area's surface. AP B.1 shows the location of the catchments, most in the North Atlantic and Baltic basins, reinforcing the need for the additional constraint for calibration described earlier in this section. AP B.1 lists some of the rivers included in each catchment.

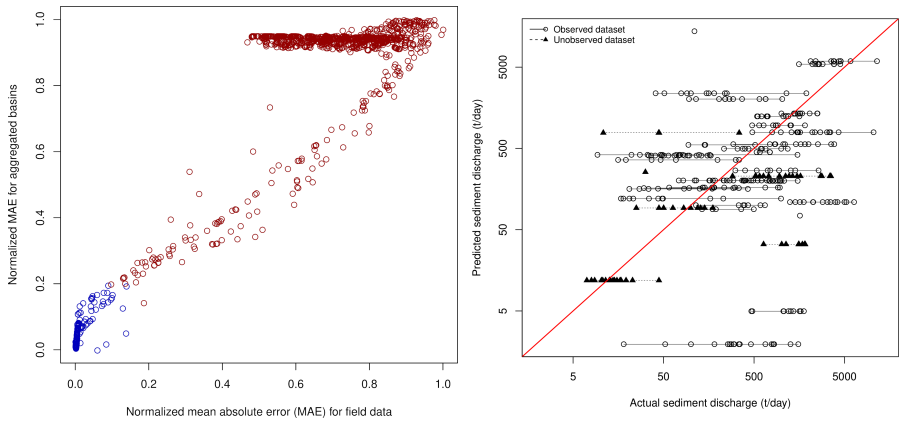


AP B.1: Location of the catchments with observations for model calibration. The red transparency is proportional to the number of observations (more transparency meaning less observations).

Model catchment	Rivers		
Baltic Sea	Ångerman	Kokemäenjoki	Skellefte
	Dalälven	Luirojoki	Stora Lule lv
	Daugava	Lule lv	Tornelven
	Elbe	Motala ström	Vistula
	Göta älv	Muonio	Vltava
	Indalsälven	Neman	Warta
	Kalix	Neris	Weser
	Kemijoki	Oder	Western Bug
	Klarälven	Oulujoki	
Black Sea	Borcea branch	Ialomița	Ráckeve-Soroksár
	Chilia branch	Inn	Danube
	Danube	Morava	Sava
	Drava	Mur	Sfântu Gheorghe
	Drina	Mureș	Sió
	Ferenc Csatorna	Olt	Sulina branch
	Great Morava	Prut	Tisza
North Atlantic Ocean	Ariège	Main	Severn
	Dordogne	Marne	Shannon
	Douro	Meuse	Tagus
	Ems	Minho	Tarn
	Esla	Moselle	Thames
	Garonne	Nederrijn	Trent
	Guadalquivir	Ness	Vienne
	Guadiana	Oich	Waal
	IJssel	Rhine	Yonne
	Lek	Scheldt	Záncara
	Loire	Seine	
Mediterranean Sea	Adige	Ebro	Rhône
	Bojana	Haliacmon	Sane
	Dora Baltea	Maritsa	Segre
	Doubs	Mincio	Struma
	Drin	Ofanto	Tiber
	Durance	Po	Ticino

AP B.1: List of some of the rivers included in each model catchment. For limitations, see Section S5.

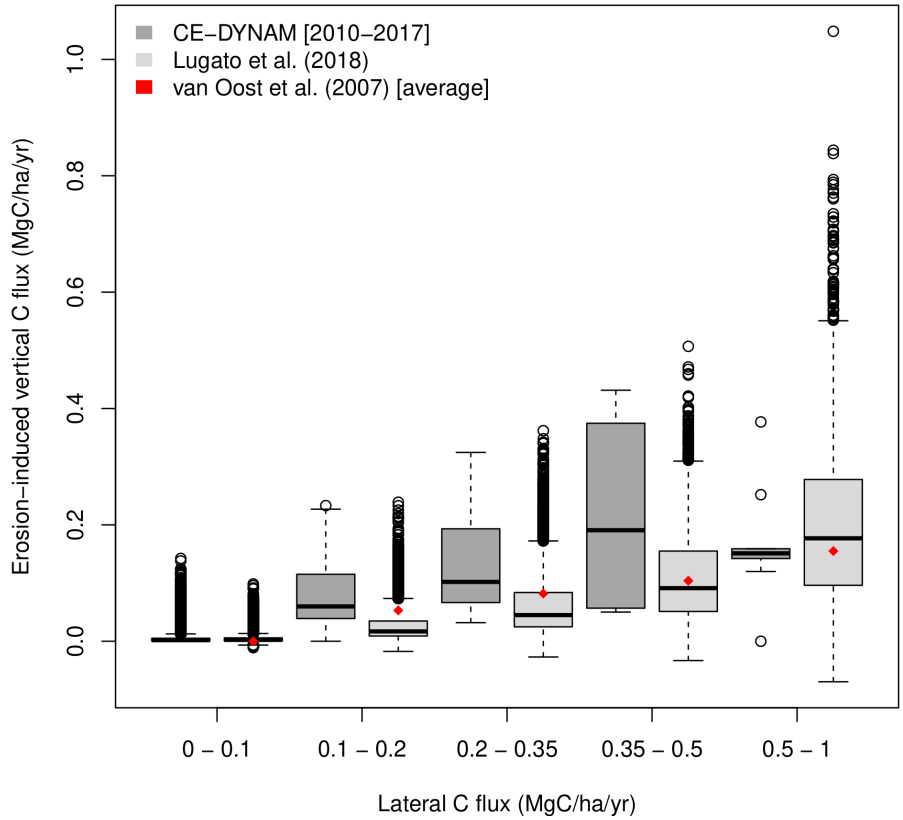
The normalized results for the 1350 random points are shown in AP B.2, left. The group of points in the bottom left (in blue) was manually inspected for presenting small errors for the sediment discharge observations and the aggregated basins data simultaneously (horizontal and vertical axis, respectively). The plot of predicted vs. observed for the best set of parameters after such inspection is shown in AP B.2, right. In this case, the Pearson correlation between datasets was calculated as t =0.41 and v =0.15 for training and validation, respectively. For the training dataset, the average observed value was 1035.49 t/day, with a median absolute error of 842.13 t/day. For the validation, the values were 610.98 t/day and 528.87 t/day for the average observed and median absolute error, respectively.



AP B.2: Calibration results: random search (left) and predictions for the best set of parameters (right). The “Observed dataset” was used for model calibration, while the “Unobserved dataset” was used for validation.

Dynamic replacement

In eroding areas, ETD induces a lateral carbon loss likely to be offset by a higher stabilization of carbon from the atmosphere by photosynthesis, leading to a process known as a dynamic replacement. Thus, understanding the intensity of the dynamic replacement representation within the model is relevant, as it can point towards the under or overestimation of the continuous carbon sink effect induced by erosion. In order to evaluate such an effect, we compared our results to the inventory made by Oost et al. (2007) and the simulation results of Lugato et al. (2018). As shown in AP B.3, our simulated results are reasonable for all classes of lateral carbon flux, presenting median values that nearly coincide with the averages reported by Oost et al. (2007), and presenting a distribution slightly shifted to higher values if compared against Lugato et al. (2018). The class of 0.35 - 0.5 MgC/ha/year seems particular for having an increased median if compared to the other works.



AP B.3: Comparison between the dynamic replacement for croplands in CE-DYNAM and other works (Lugato et al., 2018; Oost et al., 2007)

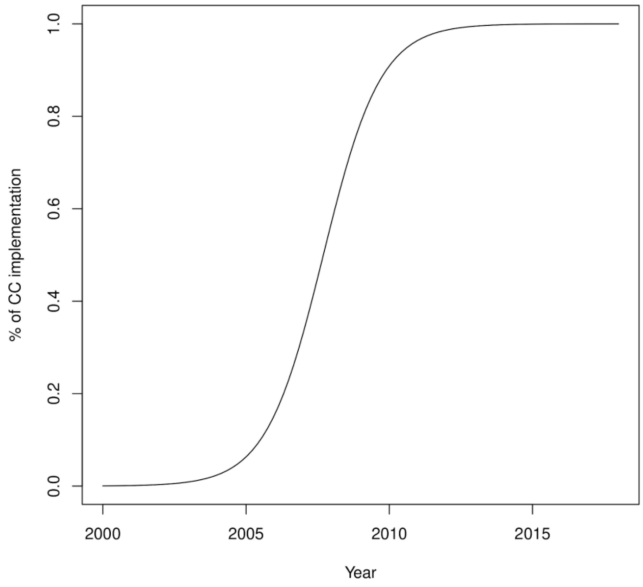
.

Assumptions for the cover crop scenarios and uncertainties

Due to the low availability of spatially-explicit data for cover crops (CCs), elaborating their application scenarios was only possible with extra assumptions. For instance, CC_Current and CC_Theoretic are defined in terms of their target CC spatial distribution. This means that, despite the known distribution around the end of simulations, nothing is said or known about the initial moment nor the rate of their implementation. Here, we tried to overcome this problem by introducing a function relating the years to the share of CC implementation, which is included in the calculation as a multiplier to the CC map of the scenario. This way, when the spatial distribution is known, the multiplier equals 100% at the target year.

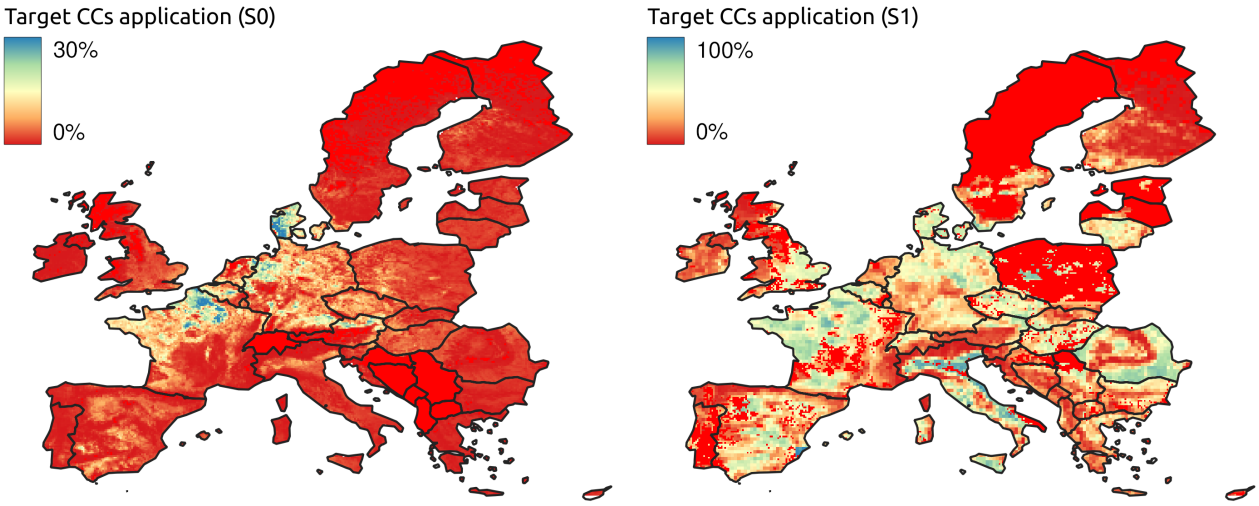
In order to construct the multiplier function, we searched for relevant clues to guide our assumed trajectory. The major clue used was the Nitrates Directive, which is known to be a driver of CC application in Europe (Kathage et al., 2022). The directive was created in 1991,

and was "severely behind schedule" in 1998 (Goodchild, 1998), but in 2000, Member States demonstrated commitment to improve its implementation (European Environment Agency, 2022). Furthermore, the Farm Structure Survey (European Commission, 2022), which contains the official statistics of farm management across Europe, indicates a stagnation of CC implementation around 2010 if compared to 2016. To combine these pieces of information, we opted to construct our assumption based on fitting a logistic curve to three reasonable points: the slow start of CC implementation around 2000, the stagnation in 2010, and the full implementation around 2016 (AP B.4).

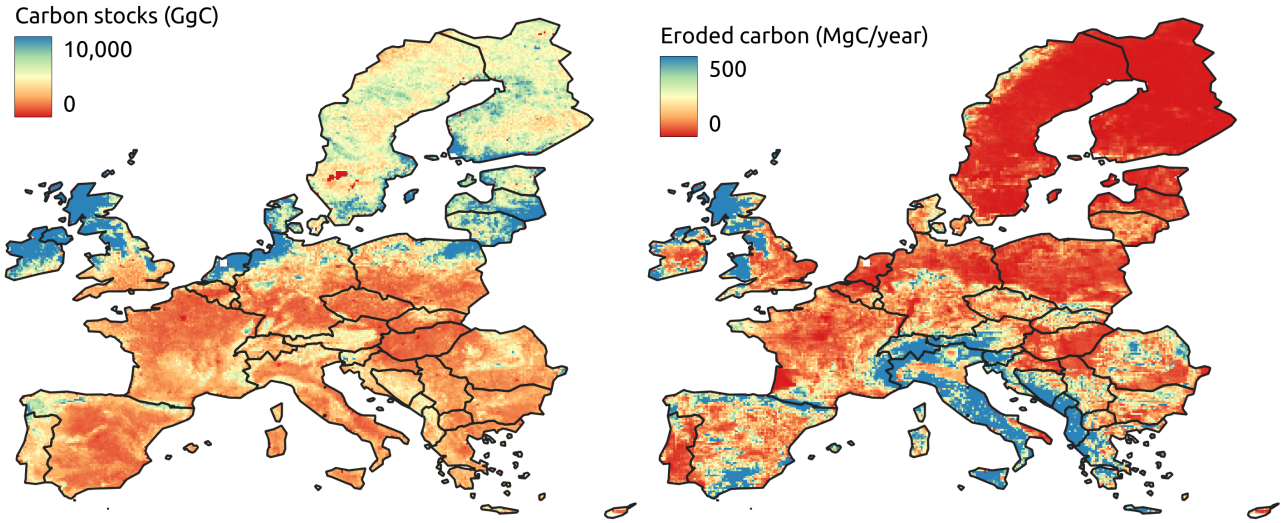


AP B.4: Assumed function relating the year to the share of CC implementation.

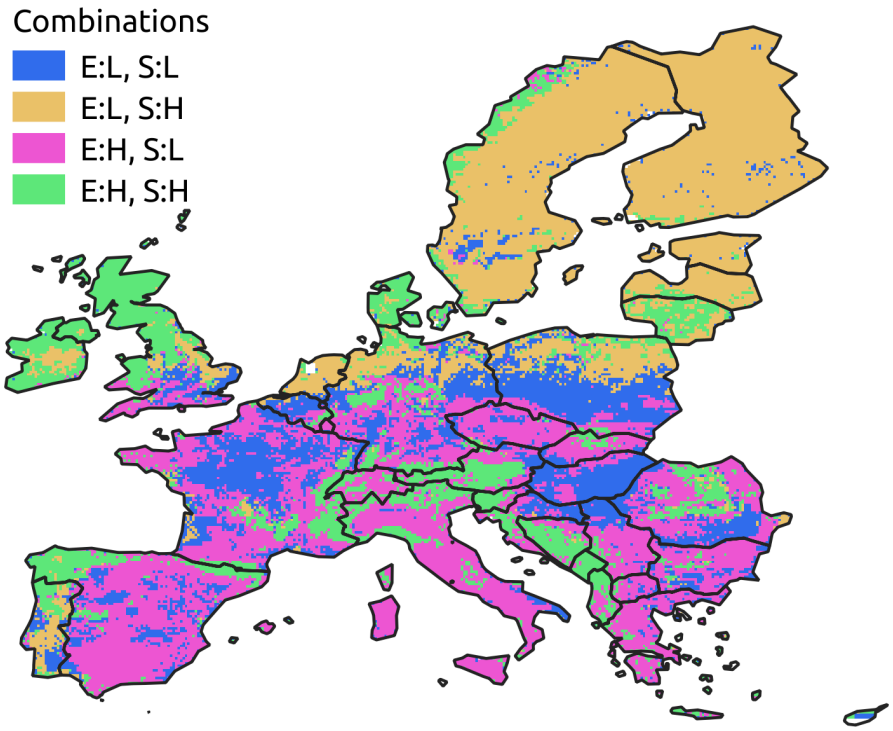
Another assumption concerned the incremental carbon input due to the adoption of CCs. Because CE-DYNAM uses only an emulator of the carbon cycle, it is impossible to represent such additional input mechanistically. In order to overcome this problem, we ran two DayCent simulations (Parton et al., 1998; Lugato et al., 2018), one without and one with CC application in the whole area of croplands in Europe. Then, we isolated the extra net primary productivity (NPP) input due to CCs by subtracting the two simulation outputs. Because the NPP input without CCs proved to be comparable between DayCent and ORCHIDEE, such isolated input was inserted directly into CE-DYNAM proportional to the CC application of each scenario. The target CC application of CC_Current and CC_Theoretic can be seen in AP B.5, left and right, respectively. For reference, AP B.6 shows the SOC stocks before erosion (left) and the corresponding eroded carbon (right) and AP B.7 shows the combinations between a reclassification of 'low' and 'high' for both.



AP B.5: Target CCs spatial distribution for CC_Current (left) and CC_Theoretic (right). The shares refer to the total land and do not exceed the shares of croplands in each pixel.



AP B.6: SOC stocks for non-eroded carbon (left) and total eroded carbon per year (right).



AP B.7: Combinations of classes with low/high erosion and low/high SOC stock. For each variable, the threshold was set to be its median. The legend format is <variable>:<reclassification>, so E:L means erosion: low and S:H means stock: high.

In CE-DYNAM, the incorporation of model uncertainties through standard methods such as Monte Carlo simulations is precluded by the long model runtimes (see section S2). Therefore, an alternative method had to be used to include the uncertainty in carbon simulations due to two of its parameters: the enrichment factor, corresponding to the carbon concentration in the sediment laterally transported, and the effect of CCs on erosion reduction. The method consisted of setting the fixed values that each parameter is allowed to assume and then simulating all possible combinations. For the enrichment factor, we set 2 possible values: 1.0, indicating no enrichment, and 2.0, based on Lugato et al. (2018). For the effect of CCs on erosion reduction, we set 3 possible values: 15%, 20%, and 23%, corresponding to the results of Nyakatawa et al. (2001), Verstraeten et al. (2006) and G. J. Wall and Shelton (2002), respectively. Consequently, a total of 15 carbon simulations were run:

$$N = \overbrace{2}^{\text{CCs scenarios}} * \overbrace{2}^{\text{Enr. factors}} * \overbrace{3}^{\text{Eros. reduction}} + \overbrace{1}^{\text{No CCs, With ETD}} * \overbrace{2}^{\text{Enr. factors}} + \overbrace{1}^{\text{No CCs, No ETD}} = 15$$

Model limitations

This section describes some of the limitations of the current approach. First, the carbon emulator only imitates land surface models, representing in an implicit way the many processes (i.e., ploughing, harvesting etc.) that are present explicitly in ORCHIDEE or DayCENT. This means that the erosion feedback on, for example, microbial respiration, is not captured in CE-DYNAM. Second, the model currently does not include the effect of erosion on the net primary productivity, which may lead to under or overestimation of effects. Third, rivers are not explicitly represented in the model, which does not allow us to calculate the transfer to rivers and dams. Even though this could be done by setting thresholds to the flow accumulation information, as done by similar models such as Watem-SEDEM, we opted for not including additional assumptions. Fourth, for the same reason, the dynamics of dissolved organic carbon (DOC) is not yet represented in CE-DYNAM. Fifth, even though large basins can be relatively well represented in a 10km spatial resolution, most of the small-scale variation can not be captured. One major limitation found about this aspect was that the Elbe river, which flows to the Atlantic Ocean, is here represented as flowing to the Baltic Sea. For all these issues, further work is necessary to increase the realism in the model representation of ETD fluxes.

Concerning the scenarios, we had to assume that the application of CCs started with their continental policies, around 2000. Even though this is not true in practice, a better assumption could not be made due to the lack of data on CCs occurrence.

Bibliography

R. Amundson, A. A. Berhe, J. W. Hopmans, C. Olson, A. E. Sztein, and D. L. Sparks. Soil and human security in the 21st century. *Science*, 348(6235), May 2015. doi: 10.1126/science.1261071. URL https://doi.org/10.1126/science.1261071.

M. M. Azam. Soil contamination and remediation measures: Revisiting the relevant laws and institutions. In *Environmental Remediation Technologies for Metal-Contaminated Soils*, pages 99–124. Springer Japan, 2016. doi: 10.1007/978-4-431-55759-3_5. URL https://doi.org/10.1007/978-4-431-55759-3_5.

Z. Bai. *Templates for the solution of algebraic eigenvalue problems : a practical guide*. Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104, Philadelphia, Pa, 2000. ISBN 9780898719581.

C. Ballabio, P. Borrelli, J. Spinoni, K. Meusburger, S. Michaelides, S. Beguería, A. Klik, S. Petan, M. Janeček, P. Olsen, J. Aalto, M. Lakatos, A. Rymszewicz, A. Dumitrescu, M. P. Tadić, N. Diodato, J. Kostalova, S. Rousseva, K. Banasik, C. Alewell, and P. Panagos. Mapping monthly rainfall erosivity in europe. *Science of The Total Environment*, 579:1298–1315, Feb. 2017. doi: 10.1016/j.scitotenv.2016.11.123. URL https://doi.org/10.1016/j.scitotenv.2016.11.123.

D. M. Bates and D. G. Watts, editors. *Nonlinear Regression Analysis and Its Applications*. John Wiley & Sons, Inc., Aug. 1988. doi: 10.1002/9780470316757. URL https://doi.org/10.1002/9780470316757.

B. Bera, S. Saha, and S. Bhattacharjee. Forest cover dynamics (1998 to 2019) and prediction of deforestation probability using binary logistic regression (BLR) model of silabati watershed, india. *Trees, Forests and People*, 2:100034, Dec. 2020. doi: 10.1016/j.tfp.2020.100034. URL https://doi.org/10.1016/j.tfp.2020.100034.

E. Beriaux, A. Jago, C. Lucau-Danila, V. Planchon, and P. Defourny. Sentinel-1 time series for crop identification in the framework of the future CAP monitoring. *Remote Sensing*, 13(14): 2785, July 2021. doi: 10.3390/rs13142785. URL https://doi.org/10.3390/rs13142785.

N. E. Blair and R. C. Aller. The fate of terrestrial organic carbon in the marine environment. *Annual Review of Marine Science*, 4(1):401–423, Jan. 2012. doi: 10.1146/annurev-marine-120709-142717. URL https://doi.org/10.1146/annurev-marine-120709-142717.

G. B. Bonan and S. C. Doney. Climate, ecosystems, and planetary futures: The challenge to predict life in earth system models. *Science*, 359(6375):eaam8328, Feb. 2018. doi: 10.1126/science.aam8328. URL https://doi.org/10.1126/science.aam8328.

H.-R. Bork and A. Lang. Quantification of past soil erosion and land use / land cover changes in germany. In *Long Term Hillslope and Fluvial System Modelling*, pages 231–239. Springer Berlin Heidelberg, 2003. doi: 10.1007/3-540-36606-7_12. URL https://doi.org/10.1007/3-540-36606-7_12.

P. Borrelli and P. Panagos. An indicator to reflect the mitigating effect of common agricultural policy on soil erosion. *Land Use Policy*, 92:104467, Mar. 2020. doi: 10.1016/j.landusepol.2020.104467. URL https://doi.org/10.1016/j.landusepol.2020.104467.

P. Borrelli, K. Paustian, P. Panagos, A. Jones, B. Schütt, and E. Lugato. Effect of good agricultural and environmental conditions on erosion and soil organic carbon balance: A national case study. *Land Use Policy*, 50:408–421, Jan. 2016. doi: 10.1016/j.landusepol.2015.09.033. URL https://doi.org/10.1016/j.landusepol.2015.09.033.

P. Borrelli, D. A. Robinson, L. R. Fleischer, E. Lugato, C. Ballabio, C. Alewell, K. Meusburger, S. Modugno, B. Schütt, V. Ferro, V. Bagarello, K. V. Oost, L. Montanarella, and P. Panagos. An assessment of the global impact of 21st century land use change on soil erosion. *Nature Communications*, 8(1), Dec. 2017. doi: 10.1038/s41467-017-02142-7. URL https://doi.org/10.1038/s41467-017-02142-7.

P. Borrelli, K. V. Oost, K. Meusburger, C. Alewell, E. Lugato, and P. Panagos. A step towards a holistic assessment of soil degradation in europe: Coupling on-site erosion with sediment transfer and carbon fluxes. *Environmental Research*, 161:291–298, Feb. 2018a. doi: 10.1016/j.envres.2017.11.009. URL https://doi.org/10.1016/j.envres.2017.11.009.

P. Borrelli, K. van Oost, K. Meusburger, C. Alewell, E. Lugato, and P. Panagos. A step towards a holistic assessment of soil degradation in europe: Coupling on-site erosion with sediment transfer and carbon fluxes. *Environmental Research*, 161:291–298, Feb. 2018b. doi: 10.1016/j.envres.2017.11.009. URL https://doi.org/10.1016/j.envres.2017.11.009.

P. Borrelli, D. A. Robinson, P. Panagos, E. Lugato, J. E. Yang, C. Alewell, D. Wuepper, L. Montanarella, and C. Ballabio. Land use and climate change impacts on global soil erosion by water (2015-2070). *Proceedings of the National Academy of Sciences*, 117(36):21994–22001, 2020. ISSN 0027-8424. doi: 10.1073/pnas.2001403117. URL https://www.pnas.org/content/117/36/21994.

P. Borrelli, C. Alewell, P. Alvarez, J. A. A. Anache, J. Baartman, C. Ballabio, N. Bezak, M. Biddoccu, A. Cerdà, D. Chalise, S. Chen, W. Chen, A. M. D. Girolamo, G. D. Gessesse, D. Deumlich, N. Diodato, N. Efthimiou, G. Erpul, P. Fiener, M. Freppaz, F. Gentile, A. Gericke, N. Haregeweyn, B. Hu, A. Jeanneau, K. Kaffas, M. Kiani-Harchegani, I. L. Villuendas, C. Li, L. Lombardo, M. López-Vicente, M. E. Lucas-Borja, M. Märker, F. Matthews, C. Miao,

M. Mikoš, S. Modugno, M. Möller, V. Naipal, M. Nearing, S. Owusu, D. Panday, E. Patault, C. V. Patriche, L. Poggio, R. Portes, L. Quijano, M. R. Rahdari, M. Renima, G. F. Ricci, J. Rodrigo-Comino, S. Saia, A. N. Samani, C. Schillaci, V. Syrris, H. S. Kim, D. N. Spinola, P. T. Oliveira, H. Teng, R. Thapa, K. Vantas, D. Vieira, J. E. Yang, S. Yin, D. A. Zema, G. Zhao, and P. Panagos. Soil erosion modelling: A global review and statistical analysis. *Science of The Total Environment*, 780:146494, Aug. 2021. doi: 10.1016/j.scitotenv.2021.146494. URL https://doi.org/10.1016/j.scitotenv.2021.146494.

P. Borrelli, C. Ballabio, J. E. Yang, D. A. Robinson, and P. Panagos. GloSEM: High-resolution global estimates of present and future soil displacement in croplands by water erosion. *Scientific Data*, 9(1), July 2022. doi: 10.1038/s41597-022-01489-x. URL https://doi.org/10.1038/s41597-022-01489-x.

J. S. Bridge. *Rivers and floodplains : forms, processes, and sedimentary record*. Blackwell Pub, Oxford, UK Malden, MA, USA, 2003. ISBN 978-0-632-06489-2.

J. L. Brooke. The holocene. In *The Palgrave Handbook of Climate History*, pages 175–182. Palgrave Macmillan UK, 2018. doi: 10.1057/978-1-137-43020-5_15. URL https://doi.org/10.1057/978-1-137-43020-5_15.

G. Büttner and B. Kosztra. Clc2018 technical guidelines. https://land.copernicus.eu/user-corner/technical-library/clc2018technicalguidelines_final.pdf, 2022. Accessed: 2022-09-21.

M. Camino-Serrano, B. Guenet, S. Luyssaert, P. Ciais, V. Bastrikov, B. D. Vos, B. Gielen, G. Gleixner, A. Jornet-Puig, K. Kaiser, D. Kothawala, R. Lauerwald, J. Peñuelas, M. Schrumpf, S. Vicca, N. Vuichard, D. Walmsley, and I. A. Janssens. ORCHIDEE-SOM: modeling soil organic carbon (SOC) and dissolved organic carbon (DOC) dynamics along vertical soil profiles in europe. *Geoscientific Model Development*, 11(3):937–957, Mar. 2018. doi: 10.5194/gmd-11-937-2018. URL https://doi.org/10.5194/gmd-11-937-2018.

CarbonBrief. Mapped: how climate change affects extreme weather around the world. https://www.carbonbrief.org/mapped-how-climate-change-affects-extreme-weather-around-the-world, 2021. Accessed: 2021-07-19.

D. Carrer, G. Pique, M. Ferlicoq, X. Ceamanos, and E. Ceschia. What is the potential of cropland albedo management in the fight against global warming? a case study based on the use of cover crops. *Environmental Research Letters*, 13(4):044030, Apr. 2018. doi: 10.1088/1748-9326/aab650. URL https://doi.org/10.1088/1748-9326/aab650.

O. Cerdan, G. Govers, Y. L. Bissonnais, K. van Oost, J. Poesen, N. Saby, A. Gobin, A. Vacca, J. Quinton, K. Auerswald, A. Klik, F. Kwaad, D. Raclot, I. Ionita, J. Rejman, S. Rousseva, T. Muxart, M. Roxo, and T. Dostal. Rates and spatial variations of soil erosion in europe: A study based on erosion plot data. *Geomorphology*, 122(1-2):167–177, Oct. 2010. doi: 10.1016/j.geomorph.2010.06.011. URL https://doi.org/10.1016/j.geomorph.2010.06.011.

J. Chang, P. Ciais, T. Gasser, P. Smith, M. Herrero, P. Havlík, M. Obersteiner, B. Guenet, D. S. Goll, W. Li, V. Naipal, S. Peng, C. Qiu, H. Tian, N. Viovy, C. Yue, and D. Zhu. Climate warming from managed grasslands cancels the cooling effect of carbon sinks in sparsely grazed and natural grasslands. *Nature Communications*, 12(1), Jan. 2021. doi: 10.1038/s41467-020-20406-7. URL https://doi.org/10.1038/s41467-020-20406-7.

A. Chappell, J. Baldock, and J. Sanderman. The global significance of omitting soil erosion from soil organic carbon cycling schemes. *Nature Climate Change*, 6(2):187–191, Oct. 2015. doi: 10.1038/nclimate2829. URL https://doi.org/10.1038/nclimate2829.

P. Ciais, C. Sabine, G. Bala, L. Bopp, V. Brovkin, et al. Carbon and other biogeochemical cycles. In *Climate Change 2013 - The Physical Science Basis*, pages 465–570. Cambridge University Press, 2013.

A. Comber and W. Zeng. Spatial interpolation using areal features: A review of methods and opportunities using new forms of data with coded illustrations. *Geography Compass*, 13(10), Aug. 2019. doi: 10.1111/gec3.12465. URL https://doi.org/10.1111/gec3.12465.

E. Commission. Commission staff working document swd(2018) 301 final. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018SC0301&from=EN, 2018. Accessed: 2022-09-24.

E. Commission. List of potential agricultural practices that eco-schemes could support. https://agriculture.ec.europa.eu/system/files/2021-01/factsheet-agri-practices-under-ecoscheme_en_0.pdf, 2021. Accessed: 2022-10-23.

E. Commission. Glossary: Farm structure survey (fss). https://web.archive.org/web/20220901054908/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Farm_structure_survey_(FSS), 2022a. Accessed: 2022-09-19.

E. Commission. Eurostat: soil cover by nuts2 regions. https://ec.europa.eu/eurostat/databrowser/view/ef_mp_soil/default/table?lang=en, 2022b. Accessed: 2022-09-20.

E. Commission. Glossary: Soil cover. https://web.archive.org/web/20220119025400/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Soil_cover, 2022c. Accessed: 2022-09-19.

E. Commission. Eurostat: Farm structure (ef). https://ec.europa.eu/eurostat/cache/metadata/EN/ef_esqrs_de.htm, 2022d. Accessed: 2022-09-22.

E. Commission. Glossary: Survey on agricultural production methods (sapm). https://web.archive.org/web/20220901012231/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Survey_on_agricultural_production_methods_(SAPM), 2022e. Accessed: 2022-09-23.

E. Commission. Eurostat: Lucas - land use and land cover survey. https://web.archive.org/web/20220901034430/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=LUCAS_-_Land_use_and_land_cover_survey, 2022f. Accessed: 2022-10-10.

E. Commission. Proposed cap strategic plans and commission observations: summary overview for 27 member states. https://agriculture.ec.europa.eu/system/files/2022-07/csp-overview-28-plans-overview-june-2022_en.pdf, 2022g. Accessed: 2022-10-22.

Copernicus. Clc 2018. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018, 2022. Accessed: 2022-09-20.

R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5(1):329–359, Dec. 1996. doi: 10.1007/bf02124750. URL https://doi.org/10.1007/bf02124750.

R. d'Andrimont, A. Verhegghen, G. Lemoine, P. Kempeneers, M. Meroni, and M. van der Velde. From parcel to continental scale – a first european crop type map based on sentinel-1 and LUCAS copernicus in-situ observations. *Remote Sensing of Environment*, 266:112708, Dec. 2021. doi: 10.1016/j.rse.2021.112708. URL https://doi.org/10.1016/j.rse.2021.112708.

A. de Roo, V. Jetten, C. Wesseling, and C. Ritsema. LISEM: A physically-based hydrologic and soil erosion catchment model. In *Modelling Soil Erosion by Water*, pages 429–440. Springer Berlin Heidelberg, 1998. doi: 10.1007/978-3-642-58913-3_32. URL https://doi.org/10.1007/978-3-642-58913-3_32.

J. A. Delgado, D. W. Reeves, and R. F. Follett. Cover crops. In *Encyclopedia of Soil Science, Third Edition*, pages 484–487. CRC Press, Jan. 2017. doi: 10.1081/e-ess3-120054071. URL https://doi.org/10.1081/e-ess3-120054071.

B. Dimassi, B. Guenet, N. P. Saby, F. Munoz, M. Bardy, F. Millet, and M. P. Martin. The impacts of CENTURY model initialization scenarios on soil organic carbon dynamics simulation in french long-term experiments. *Geoderma*, 311:25–36, Feb. 2018. doi: 10.1016/j.geoderma.2017.09.038. URL https://doi.org/10.1016/j.geoderma.2017.09.038.

S. Doetterl, A. A. Berhe, E. Nadeu, Z. Wang, M. Sommer, and P. Fiener. Erosion, deposition and soil carbon: A review of process-level controls, experimental tools and models to address c cycling in dynamic landscapes. *Earth-Science Reviews*, 154:102–122, Mar. 2016. doi: 10.1016/j.earscirev.2015.12.005. URL https://doi.org/10.1016/j.earscirev.2015.12.005.

M. Dotterweich. The history of soil erosion and fluvial deposits in small catchments of central europe: Deciphering the long-term interaction between humans and the environment — a review. *Geomorphology*, 101(1-2):192–208, Oct. 2008. doi: 10.1016/j.geomorph.2008.05.023. URL https://doi.org/10.1016/j.geomorph.2008.05.023.

D. Eddelbuettel and J. J. Balamuta. Extending *R* with *C++*: A Brief Introduction to *Rcpp*. *The American Statistician*, 72(1):28–36, 2018. doi: 10.1080/00031305.2017.1375990.

D. Eddelbuettel and C. Sanderson. Rcpparmadillo: Accelerating *R* with high-performance *C++* linear algebra. *Computational Statistics and Data Analysis*, 71:1054–1063, March 2014. URL http://dx.doi.org/10.1016/j.csda.2013.02.005.

European Commission. European green deal, 2021. URL https://ec.europa.eu/clima/eu-action/european-green-deal_en.

European Commission. Glossary: Farm structure survey (fss). https://web.archive.org/web/20220901054908/https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Farm_structure_survey_(FSS), 2022.

European Environment Agency. Nitrate directive. https://www.eea.europa.eu/archived/archived-content-water-topic/water-pollution/prevention-strategies/nitrate-directive, 2022.

European Spatial Agency. Climate change initiative. https://www.esa-landcover-cci.org/, 2021.

European Union. The eu - what it is and what it does, 2021. URL https://op.europa.eu/webpub/com/eu-what-it-is/en/.

B. S. Everitt and A. Skrondal. *The Cambridge Dictionary of Statistics*. Cambridge University Press, hardcover edition, 10 2010. ISBN 978-0521766999.

FAO. *Status of the World's Soil Resources: Main Report*. Food and Agriculture Organization of the United Nations, Rome, Italy, 2014. ISBN 9789251090046.

T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, et al. The shuttle radar topography mission. *Reviews of Geophysics*, 45(2), 2007. doi: https://doi.org/10.1029/2005RG000183. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005RG000183.

A. N. Fendrich, P. Ciais, E. Lugato, M. Carozzi, B. Guenet, P. Borrelli, V. Naipal, M. McGrath, P. Martin, and P. Panagos. Matrix representation of lateral soil movements: scaling and calibrating CE-DYNAM (v2) at a continental level. *Geoscientific Model Development*, 15(20): 7835–7857, Oct. 2022a. doi: 10.5194/gmd-15-7835-2022. URL https://doi.org/10.5194/gmd-15-7835-2022.

A. N. Fendrich, E. S. H. Neto, L. E. M. e Moreira, and D. D. Neto. A scalable method for the estimation of spatial disaggregation models. *Computers & Geosciences*, 166:105161, Sept. 2022b. doi: 10.1016/j.cageo.2022.105161. URL https://doi.org/10.1016/j.cageo.2022.105161.

A. N. Fendrich, F. Matthews, E. V. Eynde, M. Carozzi, Z. Li, R. d'Andrimont, E. Lugato, P. Martin, P. Ciais, and P. Panagos. From regional to parcel scale: A high-resolution map of cover crops across europe combining satellite data with statistical surveys. *Science of The Total Environment*, 873:162300, May 2023. doi: 10.1016/j.scitotenv.2023.162300. URL https://doi.org/10.1016/j.scitotenv.2023.162300.

X. Feng, Y. Wang, L. Chen, B. Fu, and G. Bai. Modeling soil erosion and its response to land-use change in hilly catchments of the chinese loess plateau. *Geomorphology*, 118(3-4):239–248, June 2010. doi: 10.1016/j.geomorph.2010.01.004. URL https://doi.org/10.1016/j.geomorph.2010.01.004.

B. Fosu, S. Wang, and K. Pegion. Synoptic and climate attributions of the december 2015 extreme flooding in missouri, USA. *Water*, 10(4):350, Mar. 2018. doi: 10.3390/w10040350. URL https://doi.org/10.3390/w10040350.

P. Friedlingstein. Carbon cycle feedbacks and future climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2054): 20140421, Nov. 2015. doi: 10.1098/rsta.2014.0421. URL https://doi.org/10.1098/rsta.2014.0421.

P. Friedlingstein, P. Cox, R. Betts, L. Bopp, W. von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, G. Bala, J. John, C. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H. D. Matthews, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K.-G. Schnitzler, R. Schnur, K. Strassmann, A. J. Weaver, C. Yoshikawa, and N. Zeng. Climate–carbon cycle feedback analysis: Results from the c4mip model intercomparison. *Journal of Climate*, 19(14):3337–3353, July 2006. doi: 10.1175/jcli3800.1. URL https://doi.org/10.1175/jcli3800.1.

E. A. P. G. J. Wall, D. R. Coote and I. J. Shelton. Ruslefac – revised universal soil loss equation for application in canada: A handbook for estimating soil loss from water erosion in canada, 2002. URL https://sis.agr.gc.ca/cansis/publications/manuals/2002-92/index.html.

V. Galy, B. Peucker-Ehrenbrink, and T. Eglinton. Global carbon export from the terrestrial biosphere controlled by erosion. *Nature*, 521(7551):204–207, May 2015. doi: 10.1038/nature14400. URL https://doi.org/10.1038/nature14400.

G. Garland, A. Edlinger, S. Banerjee, F. Degrune, P. García-Palacios, D. S. Pescador, C. Herzog, S. Romdhane, A. Saghai, A. Spor, C. Wagg, S. Hallin, F. T. Maestre, L. Philippot, M. C. Rillig, and M. G. A. van der Heijden. Crop cover is more important than rotational diversity for soil multifunctionality and cereal yields in european cropping systems. *Nature Food*, 2(1):28–37, Jan. 2021. doi: 10.1038/s43016-020-00210-8. URL https://doi.org/10.1038/s43016-020-00210-8.

L. V. Gatti, L. S. Basso, J. B. Miller, M. Gloor, L. G. Domingues, et al. Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595(7867):388–393, July 2021. doi: 10.1038/s41586-021-03629-6. URL https://doi.org/10.1038/s41586-021-03629-6.

A. Gettelman and R. B. Rood. *Demystifying Climate Models*. Springer Berlin Heidelberg, 2016. doi: 10.1007/978-3-662-48959-8. URL https://doi.org/10.1007/978-3-662-48959-8.

R. Goodchild. EU policies for the reduction of nitrogen in water: the example of the nitrates directive. *Environmental Pollution*, 102(1):737–740, 1998. doi: 10.1016/s0269-7491(98)80106-1. URL https://doi.org/10.1016/s0269-7491(98)80106-1.

N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27, Dec. 2017. doi: 10.1016/j.rse.2017.06.031. URL https://doi.org/10.1016/j.rse.2017.06.031.

D. Grados, K. Butterbach-Bahl, J. Chen, K. J. van Groenigen, J. E. Olesen, J. W. van Groenigen, and D. Abalos. Synthesizing the evidence of nitrous oxide mitigation practices in agroecosystems. *Environmental Research Letters*, 17(11):114024, Oct. 2022. doi: 10.1088/1748-9326/ac9b50. URL https://doi.org/10.1088/1748-9326/ac9b50.

J. L. Greathouse and M. Daga. Efficient sparse matrix-vector multiplication on GPUs using the CSR storage format. In *SC14: International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE, Nov. 2014. doi: 10.1109/sc.2014.68. URL https://doi.org/10.1109/sc.2014.68.

B. Guenet, B. Gabrielle, C. Chenu, D. Arrouays, J. Balesdent, M. Bernoux, E. Bruni, J.-P. Caliman, R. Cardinael, S. Chen, P. Ciais, D. Desbois, J. Fouche, S. Frank, C. Henault, E. Lugato, V. Naipal, T. Nesme, M. Obersteiner, S. Pellerin, D. S. Powlson, D. P. Rasse, F. Rees, J.-F. Soussana, Y. Su, H. Tian, H. Valin, and F. Zhou. Can n2o emissions offset the benefits from soil organic carbon storage? *Global Change Biology*, 27(2):237–256, Oct. 2020. doi: 10.1111/gcb.15342. URL https://doi.org/10.1111/gcb.15342.

L. Hamaoui-Laguel, R. Vautard, L. Liu, F. Solmon, N. Viovy, D. Khvorostyanov, F. Essl, I. Chuine, A. Colette, M. A. Semenov, A. Schaffhauser, J. Storkey, M. Thibaudon, and M. M. Epstein. Effects of climate change and seed dispersal on airborne ragweed pollen loads in europe. *Nature Climate Change*, 5(8):766–771, May 2015. doi: 10.1038/nclimate2652. URL https://doi.org/10.1038/nclimate2652.

Z. Han, M. T. Walter, and L. E. Drinkwater. N2o emissions from grain cropping systems: a meta-analysis of the impacts of fertilizer-based and ecologically-based nutrient management strategies. *Nutrient Cycling in Agroecosystems*, 107(3):335–355, Apr. 2017. doi: 10.1007/s10705-017-9836-z. URL https://doi.org/10.1007/s10705-017-9836-z.

A. B. Harper, T. Powell, P. M. Cox, J. House, C. Huntingford, T. M. Lenton, S. Sitch, E. Burke, S. E. Chadburn, W. J. Collins, E. Comyn-Platt, V. Daioglou, J. C. Doelman, G. Hayman, E. Robertson, D. van Vuuren, A. Wiltshire, C. P. Webber, A. Bastos, L. Boysen, P. Ciais, N. Devaraju, A. K. Jain, A. Krause, B. Poulter, and S. Shu. Land-use emissions play a critical role in land-based mitigation for paris climate targets. *Nature Communications*, 9(1), Aug. 2018. doi: 10.1038/s41467-018-05340-z. URL https://doi.org/10.1038/s41467-018-05340-z.

C. G. A. Harrison. Rates of continental erosion and mountain building. *Geologische Rundschau*, 83(2):431–447, July 1994. doi: 10.1007/bf00210556. URL https://doi.org/10.1007/bf00210556.

S. C. Herring, M. P. Hoerling, J. P. Kossin, T. C. Peterson, and P. A. Stott. Explaining extreme events of 2014 from a climate perspective. *Bulletin of the American Meteorological Society*, 96(12):S1–S172, Dec. 2015. doi: 10.1175/bams-explainingextremeevents2014.1. URL https://doi.org/10.1175/bams-explainingextremeevents2014.1.

R. J. Hijmans. *terra: Spatial Data Analysis*, 2022. URL https://CRAN.R-project.org/package=terra. R package version 1.6-17.

D. Hillel. *Introduction to environmental soil physics*. Elsevier Academic Press, Amsterdam Boston, 2004. ISBN 9780080495774.

P. R. Hobbs, K. Sayre, and R. Gupta. The role of conservation agriculture in sustainable agriculture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1491):543–555, July 2007. doi: 10.1098/rstb.2007.2169. URL https://doi.org/10.1098/rstb.2007.2169.

T. Hoffmann, S. M. Mudd, K. van Oost, G. Verstraeten, G. Erkens, A. Lang, H. Middelkoop, J. Boyle, J. O. Kaplan, J. Willenbring, and R. Aalto. Short communication: Humans and the missing c-sink: erosion and burial of soil carbon through time. *Earth Surface Dynamics*, 1(1):45–52, Nov. 2013a. doi: 10.5194/esurf-1-45-2013. URL https://doi.org/10.5194/esurf-1-45-2013.

T. Hoffmann, M. Schlummer, B. Notebaert, G. Verstraeten, and O. Korup. Carbon burial in soil sediments from holocene agricultural erosion, central europe. *Global Biogeochemical Cycles*, 27(3):828–835, Aug. 2013b. doi: 10.1002/gbc.20071. URL https://doi.org/10.1002/gbc.20071.

Y. Huang, X. Lu, Z. Shi, D. Lawrence, C. D. Koven, J. Xia, Z. Du, E. Kluzek, and Y. Luo. Matrix approach to land carbon cycle modeling: A case study with the community land model. *Global Change Biology*, 24(3):1394–1404, Nov. 2017. doi: 10.1111/gcb.13948. URL https://doi.org/10.1111/gcb.13948.

Y. Huang, D. Zhu, P. Ciais, B. Guenet, Y. Huang, D. S. Goll, M. Guimberteau, A. Jornet-Puig, X. Lu, and Y. Luo. Matrix-based sensitivity assessment of soil organic carbon storage: A case study from the orchidee-mict model. *Journal of Advances in Modeling Earth Systems*, 10(8):1790–1808, 2018. doi: https://doi.org/10.1029/2017MS001237. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017MS001237.

R. Huggett. *Fundamentals of geomorphology*. Routledge, London New York, NY, 2017. ISBN 9781138940642.

Institut National de l'Information Géographique et Forestière. Registre parcellaire graphique. https://geoservices.ign.fr/rpg, 2022. Accessed: 2022-10-06.

IPCC. Summary for policymakers. In *Climate Change 2013 - The Physical Science Basis*, pages 1–30. Cambridge University Press, 2013a. doi: 10.1017/cbo9781107415324.004. URL https://doi.org/10.1017/cbo9781107415324.004.

IPCC. Technical summary. In I. P. on Climate Change, editor, *Climate Change 2013 - The Physical Science Basis*, pages 31–116. Cambridge University Press, 2013b. doi: 10.1017/cbo9781107415324.005. URL https://doi.org/10.1017/cbo9781107415324.005.

IPCC. Summary for policymakers. In *Climate Change 2014 Impacts, Adaptation, and Vulnerability*, pages 1–32. Cambridge University Press, 2014. doi: 10.1017/cbo9781107415379.003. URL https://doi.org/10.1017/cbo9781107415379.003.

IPCC. *Climate change and land : an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems : summary for policymakers*. Intergovernmental Panel on Climate Change, Geneva, 2019. ISBN 9789291691548.

IPSL/LSCE. Verify project. https://verify.lsce.ipsl.fr/index.php/presentation, 2021.

J. Kathage, I Pérez Domínguez. Adoption of cover crops for climate mitigation in the eu. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC116730/ccc_report.pdf, 2019. Accessed: 2022-10-13.

J.-P. Jenny, S. Koirala, I. Gregory-Eaves, P. Francus, C. Niemann, B. Ahrens, V. Brovkin, A. Baud, A. E. K. Ojala, A. Normandeau, B. Zolitschka, and N. Carvalhais. Human and climate global-scale imprint on sediment transfer during the holocene. *Proceedings of the National Academy of Sciences*, 116(46):22972–22976, Oct. 2019. doi: 10.1073/pnas.1908179116. URL https://doi.org/10.1073/pnas.1908179116.

P. Jia and A. E. Gaughan. Dasymetric modeling: A hybrid approach using land cover and tax parcel data for mapping population in alachua county, florida. *Applied Geography*, 66: 100–108, Jan. 2016. doi: 10.1016/j.apgeog.2015.11.006. URL https://doi.org/10.1016/j.apgeog.2015.11.006.

J. Jian, B. J. Lester, X. Du, M. S. Reiter, and R. D. Stewart. A calculator to quantify cover crop effects on soil health and productivity. *Soil and Tillage Research*, 199:104575, May 2020. doi: 10.1016/j.still.2020.104575. URL https://doi.org/10.1016/j.still.2020.104575.

H. P. Jones, D. G. Hole, and E. S. Zavaleta. Harnessing nature to help people adapt to climate change. *Nature Climate Change*, 2(7):504–509, June 2012. doi: 10.1038/nclimate1463. URL https://doi.org/10.1038/nclimate1463.

Journal Officiel de la République Française. Arrêté du 9 avril 2018 modifiant l'arrêté du 12 novembre 2015 fixant certaines dispositions relatives au paiement pour les pratiques agricoles bénéfiques pour le climat et l'environnement dit « paiement vert » prévu par la

politique agricole commune. https://www.legifrance.gouv.fr/download/pdf?id=-TK_H2XcPnMCxk14S-Q3vZtam0AAwx52TSzzQW0k1-4=, 2018. Accessed: 2022-10-04.

JRC. Nitrated directive - vulnerable zones reporting 7. https://water.jrc.ec.europa.eu/portal/apps/webappviewer/index.html?id=b33a220c1b284583851e93a245da02ef, 2022. Accessed: 2022-10-21.

JRC. European soil data centre. https://esdac.jrc.ec.europa.eu/, 2023. Accessed: 2023-01-04.

P. Julien. *Erosion and sedimentation*. Cambridge University Press, Cambridge New York, 2010. ISBN 9780521830386.

T. Kaspar and J. Singer. The use of cover crops to manage soil. In *Soil Management: Building a Stable Base for Agriculture*, pages 321–337. Soil Science Society of America, Nov. 2015. doi: 10.2136/2011.soilmanagement.c21. URL https://doi.org/10.2136/2011.soilmanagement.c21.

J. Kathage, B. Smit, B. Janssens, W. Haagsma, and J. L. Adrados. How much is policy driving the adoption of cover crops? evidence from four EU regions. *Land Use Policy*, 116: 106016, May 2022. doi: 10.1016/j.landusepol.2022.106016. URL https://doi.org/10.1016/j.landusepol.2022.106016.

S. D. Keesstra, J. Bouma, J. Wallinga, P. Tittonell, P. Smith, A. Cerdà, L. Montanarella, J. N. Quinton, Y. Pachepsky, W. H. van der Putten, R. D. Bardgett, S. Moolenaar, G. Mol, B. Jansen, and L. O. Fresco. The significance of soils and soil science towards realization of the united nations sustainable development goals. *SOIL*, 2(2):111–128, Apr. 2016. doi: 10.5194/soil-2-111-2016. URL https://doi.org/10.5194/soil-2-111-2016.

N. Kim, M. C. Zabaloy, K. Guan, and M. B. Villamil. Do cover crops benefit soil microbiome? a meta-analysis of current research. *Soil Biology and Biochemistry*, 142:107701, Mar. 2020. doi: 10.1016/j.soilbio.2019.107701. URL https://doi.org/10.1016/j.soilbio.2019.107701.

D. A. King. ENVIRONMENT: Climate change science: Adapt, mitigate, or ignore? *Science*, 303 (5655):176–177, Jan. 2004. doi: 10.1126/science.1094329. URL https://doi.org/10.1126/science.1094329.

P. Kinnell. Event soil loss, runoff and the universal soil loss equation family of models: A review. *Journal of Hydrology*, 385(1-4):384–397, May 2010. doi: 10.1016/j.jhydrol.2010.01.024. URL https://doi.org/10.1016/j.jhydrol.2010.01.024.

K. Koudahe, S. C. Allen, and K. Djaman. Critical review of the impact of cover crops on soil properties. *International Soil and Water Conservation Research*, 10(3):343–354, Sept. 2022. doi: 10.1016/j.iswcr.2022.03.003. URL https://doi.org/10.1016/j.iswcr.2022.03.003.

G. Krinner, N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19(1), Feb. 2005. doi: 10.1029/2003gb002199. URL https://doi.org/10.1029/2003gb002199.

J. Laflen. *Pioneering soil erosion prediction : the USLE story*. WASWC, Beijing, 2003. ISBN 9749131037.

J. M. Laflen and D. C. Flanagan. The development of u. s. soil erosion prediction and modeling. *International Soil and Water Conservation Research*, 1(2):1–11, Sept. 2013. doi: 10.1016/s2095-6339(15)30034-4. URL https://doi.org/10.1016/s2095-6339(15)30034-4.

R. Lal. *Global Soil Erosion by Water and Carbon Dynamics*, page 131–142. Routledge, CRC Press, 1 edition, 1995. ISBN 156670118X.

R. Lal. Soil erosion and the global carbon budget. *Environment International*, 29(4):437–450, July 2003. doi: 10.1016/s0160-4120(02)00192-7. URL https://doi.org/10.1016/s0160-4120(02)00192-7.

R. Lal. Influence of soil erosion on carbon dynamics in the world. In *Soil Erosion and Carbon Dynamics (Advances in soil science)*, pages 23–35. CRC Press, 2005.

R. Lal. Accelerated soil erosion as a source of atmospheric CO2. *Soil and Tillage Research*, 188:35–40, May 2019. doi: 10.1016/j.still.2018.02.001. URL https://doi.org/10.1016/j.still.2018.02.001.

C. Launay, S. Houot, S. Frédéric, R. Girault, F. Levavasseur, S. Marsac, and J. Constantin. Incorporating energy cover crops for biogas production into agricultural systems: benefits and environmental impacts. a review. *Agronomy for Sustainable Development*, 42(4), June 2022. doi: 10.1007/s13593-022-00790-8. URL https://doi.org/10.1007/s13593-022-00790-8.

B. Lehner, K. Verdin, and A. Jarvis. New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10):93–94, 2008. doi: https://doi.org/10.1029/2008EO100001. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008EO100001.

G. Leng and J. Hall. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. *Science of The Total Environment*, 654:811–821, 2019. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2018.10.434. URL https://www.sciencedirect.com/science/article/pii/S0048969718343341.

F. Levavasseur, P. Martin, C. Bouty, A. Barbottin, V. Bretagnolle, O. Thérond, O. Scheurer, and N. Piskiewicz. RPG explorer: A new tool to ease the analysis of agricultural landscape dynamics with the land parcel identification system. *Computers and Electronics in Agriculture*, 127:541–552, Sept. 2016. doi: 10.1016/j.compag.2016.07.015. URL https://doi.org/10.1016/j.compag.2016.07.015.

Y. Li, Q. W. Zhang, D. C. Reicosky, M. J. Lindstrom, L. Y. Bai, and L. Li. Changes in soil organic carbon induced by tillage and water erosion on a steep cultivated hillslope in the chinese loess plateau from 1898–1954 and 1954–1998. *Journal of Geophysical Research*, 112(G1), Mar. 2007. doi: 10.1029/2005jg000107. URL https://doi.org/10.1029/2005jg000107.

D. Lorey. *Global environmental challenges of the twenty-first century : resources, consumption, and sustainable solutions*. SR Books, Wilmington, Del, 2003. ISBN 9780742581838.

LSCE. ORCHIDEE Development Section. Land Cover - Description. https://orchidas.lsce.ipsl.fr/dev/lccci/, 2021. Online; accessed 15 September 2021.

X. Lu, Z. Du, Y. Huang, D. Lawrence, E. Kluzek, N. Collier, D. Lombardozzi, N. Sobhani, E. A. G. Schuur, and Y. Luo. Full implementation of matrix approach to biogeochemistry module of clm5. *Journal of Advances in Modeling Earth Systems*, 12(11):e2020MS002105, 2020. doi: https://doi.org/10.1029/2020MS002105. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002105. e2020MS002105 2020MS002105.

E. Lugato, K. Paustian, P. Panagos, A. Jones, and P. Borrelli. Quantifying the erosion effect on current carbon budget of european agricultural soils at high spatial resolution. *Global Change Biology*, 22(5):1976–1984, 2016. doi: https://doi.org/10.1111/gcb.13198. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.13198.

E. Lugato, P. Smith, P. Borrelli, P. Panagos, C. Ballabio, et al. Soil erosion is unlikely to drive a future carbon sink in europe. *Science Advances*, 4(11):eaau3523, Nov. 2018. doi: 10.1126/sciadv.aau3523. URL https://doi.org/10.1126/sciadv.aau3523.

Y. Luo, Z. Shi, X. Lu, J. Xia, J. Liang, J. Jiang, Y. Wang, M. J. Smith, L. Jiang, A. Ahlström, B. Chen, O. Hararuk, A. Hastings, F. Hoffman, B. Medlyn, S. Niu, M. Rasmussen, K. Todd-Brown, and Y.-P. Wang. Transient dynamics of terrestrial carbon storage: mathematical foundation and its applications. *Biogeosciences*, 14(1):145–161, 2017. doi: 10.5194/bg-14-145-2017. URL https://bg.copernicus.org/articles/14/145/2017/.

C. Ma, K. Johansen, and M. F. McCabe. Monitoring irrigation events and crop dynamics using sentinel-1 and sentinel-2 time series. *Remote Sensing*, 14(5):1205, Mar. 2022. doi: 10.3390/rs14051205. URL https://doi.org/10.3390/rs14051205.

O. Maimon and L. Rokach, editors. *Data Mining and Knowledge Discovery Handbook*. Springer US, 2010. doi: 10.1007/978-0-387-09823-4. URL https://doi.org/10.1007/978-0-387-09823-4.

E. Martin. Cover crops and water quality. *Environmental Modeling and Assessment*, 24(6):605–623, Mar. 2019. doi: 10.1007/s10666-019-09657-x. URL https://doi.org/10.1007/s10666-019-09657-x.

F. Matthews, G. Verstraeten, P. Borrelli, and P. Panagos. A field parcel-oriented approach to evaluate the crop cover-management factor and time-distributed erosion risk in europe.

International Soil and Water Conservation Research, March 2023. doi: 10.1016/j.iswcr.2022.09.005. URL https://doi.org/10.1016/j.iswcr.2022.09.005.

A. K. Maurya, N. Bhargava, and D. Singh. Efficient selection of SAR features using ML based algorithms for accurate FVC estimation. *Advances in Space Research*, 70(7):1795–1809, Oct. 2022. doi: 10.1016/j.asr.2022.06.039. URL https://doi.org/10.1016/j.asr.2022.06.039.

S. C. McClelland, K. Paustian, and M. E. Schipanski. Management of cover crops in temperate climates influences soil organic carbon stocks: a meta-analysis. *Ecological Applications*, 31 (3), Mar. 2021. doi: 10.1002/eap.2278. URL https://doi.org/10.1002/eap.2278.

M. D. McDaniel, L. K. Tiemann, and A. S. Grandy. Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? a meta-analysis. *Ecological Applications*, 24(3):560–570, Apr. 2014. doi: 10.1890/13-0616.1. URL https://doi.org/10.1890/13-0616.1.

M. Meroni, R. d'Andrimont, A. Vrieling, D. Fasbender, G. Lemoine, F. Rembold, L. Seguini, and A. Verhegghen. Comparing land surface phenology of major european crops as derived from SAR and multispectral data of sentinel-1 and -2. *Remote Sensing of Environment*, 253: 112232, Feb. 2021. doi: 10.1016/j.rse.2020.112232. URL https://doi.org/10.1016/j.rse.2020.112232.

H. Metzler, Q. Zhu, W. Riley, A. Hoyt, M. Müller, and C. A. Sierra. Mathematical reconstruction of land carbon models from their numerical output: Computing soil radiocarbon from c dynamics. *Journal of Advances in Modeling Earth Systems*, 12(1), Jan. 2020. doi: 10.1029/2019ms001776. URL https://doi.org/10.1029/2019ms001776.

K. Miettinen. *Nonlinear Multiobjective Optimization*. Springer US, 1998. doi: 10.1007/978-1-4615-5563-6. URL https://doi.org/10.1007/978-1-4615-5563-6.

B. Minasny, B. P. Malone, A. B. McBratney, D. A. Angers, D. Arrouays, A. Chambers, V. Chaplot, Z.-S. Chen, K. Cheng, B. S. Das, D. J. Field, A. Gimona, C. B. Hedley, S. Y. Hong, B. Mandal, B. P. Marchant, M. Martin, B. G. McConkey, V. L. Mulder, S. O'Rourke, A. C. R. de Forges, I. Odeh, J. Padarian, K. Paustian, G. Pan, L. Poggio, I. Savin, V. Stolbovoy, U. Stockmann, Y. Sulaeman, C.-C. Tsui, T.-G. Vågen, B. van Wesemael, and L. Winowiecki. Soil carbon 4 per mille. *Geoderma*, 292:59–86, Apr. 2017. doi: 10.1016/j.geoderma.2017.01.002. URL https://doi.org/10.1016/j.geoderma.2017.01.002.

Ministère de l'Agriculture te de l'Alimentation. Cultures et précisions. liste des cultures à utiliser pour renseigner le descriptif des parcelles. https://www.telepac.agriculture.gouv.fr/telepac/pdf/tas/2020/Dossier-PAC-2020_notice_cultures-precisions.pdf, 2022. Accessed: 2022-09-20.

A. B. Møller, B. Malone, N. P. Odgers, A. Beucher, B. V. Iversen, M. H. Greve, and B. Minasny. Improved disaggregation of conventional soil maps. *Geoderma*, 341:148–160, May 2019.

doi: 10.1016/j.geoderma.2019.01.038. URL https://doi.org/10.1016/j.geoderma.2019.01.038.

L. Montanarella and P. Panagos. The relevance of sustainable soil management within the european green deal. *Land Use Policy*, 100:104950, Jan. 2021. doi: 10.1016/j.landusepol.2020.104950. URL https://doi.org/10.1016/j.landusepol.2020.104950.

D. R. Montgomery. Soil erosion and agricultural sustainability. *Proceedings of the National Academy of Sciences*, 104(33):13268–13272, Aug. 2007. doi: 10.1073/pnas.0611508104. URL https://doi.org/10.1073/pnas.0611508104.

I. Muhammad, J. Wang, U. M. Sainju, S. Zhang, F. Zhao, and A. Khan. Cover cropping enhances soil microbial biomass and affects microbial community structure: A meta-analysis. *Geoderma*, 381:114696, Jan. 2021. doi: 10.1016/j.geoderma.2020.114696. URL https://doi.org/10.1016/j.geoderma.2020.114696.

E. Nadeu, A. Gobin, P. Fiener, B. van Wesemael, and K. van Oost. Modelling the impact of agricultural management on soil carbon stocks at the regional scale: the role of lateral fluxes. *Global Change Biology*, 21(8):3181–3192, Apr. 2015. doi: 10.1111/gcb.12889. URL https://doi.org/10.1111/gcb.12889.

V. Naipal, C. Reick, J. Pongratz, and K. van Oost. Improving the global applicability of the RUSLE model – adjustment of the topographical and rainfall erosivity factors. *Geoscientific Model Development*, 8(9):2893–2913, Sept. 2015. doi: 10.5194/gmd-8-2893-2015. URL https://doi.org/10.5194/gmd-8-2893-2015.

V. Naipal, C. Reick, K. van Oost, T. Hoffmann, and J. Pongratz. Modeling long-term, large-scale sediment storage using a simple sediment budget approach. *Earth Surface Dynamics*, 4(2):407–423, May 2016. doi: 10.5194/esurf-4-407-2016. URL https://doi.org/10.5194/esurf-4-407-2016.

V. Naipal, R. Lauerwald, P. Ciais, B. Guenet, and Y. Wang. Ce-dynam (v1): a spatially explicit process-based carbon erosion scheme for use in earth system models. *Geoscientific Model Development*, 13(3):1201–1222, 2020. doi: 10.5194/gmd-13-1201-2020. URL https://gmd.copernicus.org/articles/13/1201/2020/.

J. Nash and J. Sutcliffe. River flow forecasting through conceptual models part i — a discussion of principles. *Journal of Hydrology*, 10(3):282–290, Apr. 1970. doi: 10.1016/0022-1694(70)90255-6. URL https://doi.org/10.1016/0022-1694(70)90255-6.

R. Nisbet, J. F. Elder, and G. Miner. *Handbook of statistical analysis and data mining applications*. Academic Press/Elsevier, Amsterdam ; Boston, 2009. ISBN 9780123747655 9780123750860. OCLC: ocn316327105.

C. A. Nobre, G. Sampaio, L. S. Borma, J. C. Castilla-Rubio, J. S. Silva, et al. Land-use and climate change risks in the amazon and the need of a novel sustainable development paradigm.

Proceedings of the National Academy of Sciences, 113(39):10759–10768, Sept. 2016. doi: 10.1073/pnas.1605516113. URL https://doi.org/10.1073/pnas.1605516113.

B. Nowak, G. Marliac, and A. Michaud. Estimation of winter soil cover by vegetation before spring-sown crops for mainland france using multispectral satellite imagery. *Environmental Research Letters*, 16(6):064024, May 2021. doi: 10.1088/1748-9326/ac007c. URL https://doi.org/10.1088/1748-9326/ac007c.

E. Nyakatawa, K. Reddy, and J. Lemunyon. Predicting soil erosion in conservation tillage cotton production systems using the revised universal soil loss equation (RUSLE). *Soil and Tillage Research*, 57(4):213–224, Jan. 2001. doi: 10.1016/s0167-1987(00)00178-1. URL https://doi.org/10.1016/s0167-1987(00)00178-1.

O. J. of the European Union. Council directive 91/676/eec of 12 december 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31991L0676&from=EN, 1991. Accessed: 2022-09-23.

O. J. of the European Union. Regulation (eu) 2021/2116 of the european parliament and of the council of 2 december 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R2116&from=EN, 2021. Accessed: 2022-09-23.

C. Ols, J.-C. Hervé, and J.-D. Bontemps. Recent growth trends of conifers across western europe are controlled by thermal and water constraints and favored by forest heterogeneity. *Science of The Total Environment*, 742:140453, Nov. 2020. doi: 10.1016/j.scitotenv.2020.140453. URL https://doi.org/10.1016/j.scitotenv.2020.140453.

K. Olson, S. A. Ebelhar, and J. M. Lang. Long-term effects of cover crops on crop yields, soil organic carbon stocks and sequestration. *Open Journal of Soil Science*, 04(08):284–292, 2014. doi: 10.4236/ojss.2014.48030. URL https://doi.org/10.4236/ojss.2014.48030.

K. R. Olson, M. Al-Kaisi, R. Lal, and L. Cihacek. Impact of soil erosion on soil organic carbon stocks. *Journal of Soil and Water Conservation*, 71(3):61A–67A, May 2016. doi: 10.2489/jswc.71.3.61a. URL https://doi.org/10.2489/jswc.71.3.61a.

K. V. Oost, T. A. Quine, G. Govers, S. D. Gryze, J. Six, J. W. Harden, J. C. Ritchie, G. W. McCarty, G. Heckrath, C. Kosmas, J. V. Giraldez, J. R. M. da Silva, and R. Merckx. The impact of agricultural soil erosion on the global carbon cycle. *Science*, 318(5850):626–629, Oct. 2007. doi: 10.1126/science.1145724. URL https://doi.org/10.1126/science.1145724.

A. Orgiazzi, C. Ballabio, P. Panagos, A. Jones, and O. Fernández-Ugalde. LUCAS soil, the largest expandable soil dataset for europe: a review. *European Journal of Soil Science*, 69(1):140–153, Nov. 2017. doi: 10.1111/ejss.12499. URL https://doi.org/10.1111/ejss.12499.

W. R. Osterkamp, C. R. Hupp, and M. Stoffel. The interactions between vegetation and erosion: new directions for research at the interface of ecology and geomorphology.

Earth Surface Processes and Landforms, 37(1):23–36, July 2011. doi: 10.1002/esp.2173. URL https://doi.org/10.1002/esp.2173.

C. Palm, H. Blanco-Canqui, F. DeClerck, L. Gatere, and P. Grace. Conservation agriculture and ecosystem services: An overview. *Agriculture, Ecosystems & Environment*, 187:87–105, Apr. 2014. doi: 10.1016/j.agee.2013.10.010. URL https://doi.org/10.1016/j.agee.2013.10.010.

P. Panagos, K. Meusburger, C. Ballabio, P. Borrelli, and C. Alewell. Soil erodibility in europe: A high-resolution dataset based on LUCAS. *Science of The Total Environment*, 479-480:189–200, May 2014. doi: 10.1016/j.scitotenv.2014.02.010. URL https://doi.org/10.1016/j.scitotenv.2014.02.010.

P. Panagos, C. Ballabio, P. Borrelli, K. Meusburger, A. Klik, S. Rousseva, M. P. Tadić, S. Michaelides, M. Hrabalíková, P. Olsen, J. Aalto, M. Lakatos, A. Rymszewicz, A. Dumitrescu, S. Beguería, and C. Alewell. Rainfall erosivity in europe. *Science of The Total Environment*, 511: 801–814, Apr. 2015a. doi: 10.1016/j.scitotenv.2015.01.008. URL https://doi.org/10.1016/j.scitotenv.2015.01.008.

P. Panagos, P. Borrelli, and K. Meusburger. A new european slope length and steepness factor (LS-factor) for modeling soil erosion by water. *Geosciences*, 5(2):117–126, Apr. 2015b. doi: 10.3390/geosciences5020117. URL https://doi.org/10.3390/geosciences5020117.

P. Panagos, P. Borrelli, K. Meusburger, C. Alewell, E. Lugato, and L. Montanarella. Estimating the soil erosion cover-management factor at the european scale. *Land Use Policy*, 48:38–50, Nov. 2015c. doi: 10.1016/j.landusepol.2015.05.021. URL https://doi.org/10.1016/j.landusepol.2015.05.021.

P. Panagos, P. Borrelli, K. Meusburger, E. H. van der Zanden, J. Poesen, and C. Alewell. Modelling the effect of support practices (p-factor) on the reduction of soil erosion by water at european scale. *Environmental Science & Policy*, 51:23–34, Aug. 2015d. doi: 10.1016/j.envsci.2015.03.012. URL https://doi.org/10.1016/j.envsci.2015.03.012.

P. Panagos, P. Borrelli, J. Poesen, C. Ballabio, E. Lugato, K. Meusburger, L. Montanarella, and C. Alewell. The new assessment of soil loss by water erosion in europe. *Environmental Science & Policy*, 54:438–447, Dec. 2015e. doi: 10.1016/j.envsci.2015.08.012. URL https://doi.org/10.1016/j.envsci.2015.08.012.

P. Panagos, P. Borrelli, J. Poesen, K. Meusburger, C. Ballabio, E. Lugato, L. Montanarella, and C. Alewell. Reply to "the new assessment of soil loss by water erosion in europe. panagos p. et al., 2015 environ. sci. policy 54, 438-447 - a response" by evans and boardman [environ. sci. policy 58, 11-15]. *Environmental Science & Policy*, 59:53–57, May 2016. doi: 10.1016/j.envsci.2016.02.010. URL https://doi.org/10.1016/j.envsci.2016.02.010.

P. Panagos, C. Ballabio, K. Meusburger, J. Spinoni, C. Alewell, and P. Borrelli. Towards estimates of future rainfall erosivity in europe based on REDES and WorldClim datasets. *Journal of Hydrology*, 548:251–262, May 2017. doi: 10.1016/j.jhydrol.2017.03.006. URL https://doi.org/10.1016/j.jhydrol.2017.03.006.

P. Panagos, C. Ballabio, J. Poesen, E. Lugato, S. Scarpa, L. Montanarella, and P. Borrelli. A soil erosion indicator for supporting agricultural, environmental and climate policies in the european union. *Remote Sensing*, 12(9):1365, Apr. 2020. doi: 10.3390/rs12091365. URL https://doi.org/10.3390/rs12091365.

P. Panagos, C. Ballabio, M. Himics, S. Scarpa, F. Matthews, M. Bogonos, J. Poesen, and P. Borrelli. Projections of soil loss by water erosion in europe by 2050. *Environmental Science and Policy*, 124:380–392, Oct. 2021. doi: 10.1016/j.envsci.2021.07.012. URL https://doi.org/10.1016/j.envsci.2021.07.012.

P. Panagos, M. V. Liedekerke, P. Borrelli, J. Köninger, C. Ballabio, A. Orgiazzi, E. Lugato, L. Liakos, J. Hervas, A. Jones, and L. Montanarella. European soil data centre 2.0: Soil data and knowledge in support of the eu policies. *European Journal of Soil Science*, 73(6), Nov. 2022a. doi: 10.1111/ejss.13315. URL https://doi.org/10.1111/ejss.13315.

P. Panagos, L. Montanarella, M. Barbero, A. Schneegans, L. Aguglia, and A. Jones. Soil priorities in the european union. *Geoderma Regional*, 29:e00510, June 2022b. doi: 10.1016/j.geodrs.2022.e00510. URL https://doi.org/10.1016/j.geodrs.2022.e00510.

M. Panahi, K. Khosravi, A. Golkarian, M. Roostaei, R. Barzegar, E. Omidvar, F. Rezaie, P. M. Saco, A. Sharifi, C. Jun, S. M. Bateni, C.-W. Lee, and S. Lee. A country-wide assessment of iran's land subsidence susceptibility using satellite-based InSAR and machine learning. *Geocarto International*, pages 1–23, June 2022. doi: 10.1080/10106049.2022.2086631. URL https://doi.org/10.1080/10106049.2022.2086631.

W. J. Parton, D. W. Anderson, C. V. Cole, and J. W. B. Stewart. Simulation of soil organic matter formation and mineralization in semiarid agroecosystems. In *Nutrient cycling in agricultural ecosystems*, Special Publ. No. 23. The Univ. of Georgia, College of Agriculture Experiment Stations, 1983.

W. J. Parton, J. W. B. Stewart, and C. V. Cole. Dynamics of c, n, p and s in grassland soils: a model. *Biogeochemistry*, 5(1):109–131, Feb. 1988. doi: 10.1007/bf02180320. URL https://doi.org/10.1007/bf02180320.

W. J. Parton, M. Hartman, D. Ojima, and D. Schimel. DAYCENT and its land surface submodel: description and testing. *Global and Planetary Change*, 19(1-4):35–48, Dec. 1998. doi: 10.1016/s0921-8181(98)00040-x. URL https://doi.org/10.1016/s0921-8181(98)00040-x.

G. Pe'er, Y. Zinngrebe, J. Hauck, S. Schindler, A. Dittrich, S. Zingg, T. Tscharntke, R. Oppermann, L. M. Sutcliffe, C. Sirami, J. Schmidt, C. Hoyer, C. Schleyer, and S. Lakner. Adding some green

to the greening: Improving the EU's ecological focus areas for biodiversity and farmers. *Conservation Letters*, 10(5):517–530, Jan. 2017. doi: 10.1111/conl.12333. URL https://doi.org/10.1111/conl.12333.

J. Pelletier, P. Broxton, P. Hazenberg, X. Zeng, P. Troch, et al. Global 1-km gridded thickness of soil, regolith, and sedimentary deposit layers, 2016. URL http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1304.

C. Poeplau and A. Don. Carbon sequestration in agricultural soils via cultivation of cover crops – a meta-analysis. *Agriculture, Ecosystems and Environment*, 200:33–41, Feb. 2015. doi: 10.1016/j.agee.2014.10.024. URL https://doi.org/10.1016/j.agee.2014.10.024.

J. Poesen. Soil erosion in the anthropocene: Research needs. *Earth Surface Processes and Landforms*, 43(1):64–84, Oct. 2017. doi: 10.1002/esp.4250. URL https://doi.org/10.1002/esp.4250.

L. Poggio, L. M. de Sousa, N. H. Batjes, G. B. M. Heuvelink, B. Kempen, et al. Soilgrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, 7(1):217–240, 2021. doi: 10.5194/soil-7-217-2021. URL https://soil.copernicus.org/articles/7/217/2021/.

M. J. D. Powell. The NEWUOA software for unconstrained optimization without derivatives. In *Nonconvex Optimization and Its Applications*, pages 255–297. Springer US, 2006. doi: 10.1007/0-387-30065-1_16. URL https://doi.org/10.1007/0-387-30065-1_16.

M. J. D. Powell. Developments of NEWUOA for minimization without derivatives. *IMA Journal of Numerical Analysis*, 28(4):649–664, Feb. 2008. doi: 10.1093/imanum/drm047. URL https://doi.org/10.1093/imanum/drm047.

G. Prokop. The state of EU soil policy and soil related research. *Reviews in Environmental Science and Bio/Technology*, 4(3):81–86, Aug. 2005. doi: 10.1007/s11157-005-2239-7. URL https://doi.org/10.1007/s11157-005-2239-7.

L. Quijano, S. Beguería, L. Gaspar, and A. Navas. Estimating erosion rates using 137cs measurements and WATEM/SEDEM in a mediterranean cultivated field. *CATENA*, 138:38–51, Mar. 2016. doi: 10.1016/j.catena.2015.11.009. URL https://doi.org/10.1016/j.catena.2015.11.009.

T. A. Quine and K. van Oost. Insights into the future of soil erosion. *Proceedings of the National Academy of Sciences*, 117(38):23205–23207, Sept. 2020. doi: 10.1073/pnas.2017314117. URL https://doi.org/10.1073/pnas.2017314117.

P. Quinn, K. Beven, P. Chevallier, and O. Planchon. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, 5(1):59–79, Jan. 1991. doi: 10.1002/hyp.3360050106. URL https://doi.org/10.1002/hyp.3360050106.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL https://www.R-project.org/.

K. Renard. *Predicting soil erosion by water : a guide to conservation planning with the revised universal soil loss equation (RUSLE*. U.S. Dept. of Agriculture, Agricultural Research Service For sale by the U.S. G.P.O., Supt. of Docs, Washington, D.C, 1997. ISBN 0160489385.

K. G. Renard, G. R. Foster, G. A. Weesies, and J. P. Porter. Rusle: Revised universal soil loss equation. *Journal of Soil and Water Conservation*, 46(1):30–33, 1991. ISSN 0022-4561. URL https://www.jswconline.org/content/46/1/30.

C. Riggers, C. Poeplau, A. Don, C. Frühauf, and R. Dechow. How much carbon input is required to preserve or increase projected soil organic carbon stocks in german croplands under climate change? *Plant and Soil*, 460(1-2):417–433, Jan. 2021. doi: 10.1007/s11104-020-04806-8. URL https://doi.org/10.1007/s11104-020-04806-8.

M. Robert. *Soil carbon sequestration for improved land management*. Food and Agricultural Organization of the United Nations, Rome, 2001. ISBN 9251046905.

M. Robert. Global change and carbon cycle: the position of soils and agriculture. In *Soil Erosion and Carbon Dynamics (Advances in soil science)*, pages 3–12. CRC Press, 2005.

S. Roe, C. Streck, M. Obersteiner, S. Frank, B. Griscom, L. Drouet, O. Fricko, M. Gusti, N. Harris, T. Hasegawa, Z. Hausfather, P. Havlík, J. House, G.-J. Nabuurs, A. Popp, M. J. S. Sánchez, J. Sanderman, P. Smith, E. Stehfest, and D. Lawrence. Contribution of the land sector to a 1.5 °c world. *Nature Climate Change*, 9(11):817–828, Oct. 2019. doi: 10.1038/s41558-019-0591-9. URL https://doi.org/10.1038/s41558-019-0591-9.

A. J. J. V. Rompaey, G. Verstraeten, K. van Oost, G. Govers, and J. Poesen. Modelling mean annual sediment yield using a distributed approach. *Earth Surface Processes and Landforms*, 26(11):1221–1236, Sept. 2001. doi: 10.1002/esp.275. URL https://doi.org/10.1002/esp.275.

A. V. Rompaey, P. Bazzoffi, R. J. Jones, and L. Montanarella. Modeling sediment yields in italian catchments. *Geomorphology*, 65(1-2):157–169, Feb. 2005. doi: 10.1016/j.geomorph.2004.08.006. URL https://doi.org/10.1016/j.geomorph.2004.08.006.

D. Rossell, O. Abril, and A. Bhattacharya. Approximate laplace approximations for scalable model selection. *J. R. Stat. Soc. Series B Stat. Methodol.*, 83(4):853–879, Sept. 2021.

P. Roy, R. Chakrabortty, I. Chowdhuri, S. Malik, B. Das, and S. C. Pal. Development of different machine learning ensemble classifier for gully erosion susceptibility in gandheswari watershed of west bengal, india. In *Machine Learning for Intelligent Decision Science*, pages 1–26. Springer Singapore, 2020. doi: 10.1007/978-981-15-3689-2_1. URL https://doi.org/10.1007/978-981-15-3689-2_1.

S. J. Ruis and H. Blanco-Canqui. Cover crops could offset crop residue removal effects on soil carbon and other properties: A review. *Agronomy Journal*, 109(5):1785–1805, Sept. 2017. doi: 10.2134/agronj2016.12.0735. URL https://doi.org/10.2134/agronj2016.12.0735.

J. Sanderman, T. Hengl, and G. J. Fiske. Soil carbon debt of 12, 000 years of human land use. *Proceedings of the National Academy of Sciences*, 114(36):9575–9580, Aug. 2017. doi: 10.1073/pnas.1706103114. URL https://doi.org/10.1073/pnas.1706103114.

A. Scavo, S. Fontanazza, A. Restuccia, G. R. Pesce, C. Abbate, and G. Mauromicale. The role of cover crops in improving soil fertility and plant nutritional status in temperate climates. a review. *Agronomy for Sustainable Development*, 42(5), Sept. 2022. doi: 10.1007/s13593-022-00825-0. URL https://doi.org/10.1007/s13593-022-00825-0.

J. Schewe, S. N. Gosling, C. Reyer, F. Zhao, P. Ciais, J. Elliott, L. Francois, V. Huber, H. K. Lotze, S. I. Seneviratne, M. T. H. van Vliet, R. Vautard, Y. Wada, L. Breuer, M. Büchner, D. A. Carozza, J. Chang, M. Coll, D. Deryng, A. de Wit, T. D. Eddy, C. Folberth, K. Frieler, A. D. Friend, D. Gerten, L. Gudmundsson, N. Hanasaki, A. Ito, N. Khabarov, H. Kim, P. Lawrence, C. Morfopoulos, C. Müller, H. M. Schmied, R. Orth, S. Ostberg, Y. Pokhrel, T. A. M. Pugh, G. Sakurai, Y. Satoh, E. Schmid, T. Stacke, J. Steenbeek, J. Steinkamp, Q. Tang, H. Tian, D. P. Tittensor, J. Volkholz, X. Wang, and L. Warszawski. State-of-the-art global models underestimate impacts from climate extremes. *Nature Communications*, 10(1), Mar. 2019. doi: 10.1038/s41467-019-08745-6. URL https://doi.org/10.1038/s41467-019-08745-6.

M. Schneider, A. Broszeit, and M. Körner. Eurocrops: A pan-european dataset for time series crop type classification. 2021. doi: 10.2760/125905.

G. E. Shackelford, R. Kelsey, and L. V. Dicks. Effects of cover crops on multiple ecosystem services: Ten meta-analyses of data from arable farmland in california and the mediterranean. *Land Use Policy*, 88:104204, Nov. 2019. doi: 10.1016/j.landusepol.2019.104204. URL https://doi.org/10.1016/j.landusepol.2019.104204.

W. Shangguan, Y. Dai, Q. Duan, B. Liu, and H. Yuan. A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, 6(1):249–263, 2014. doi: https://doi.org/10.1002/2013MS000293. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013MS000293.

C. A. Sierra and M. Müller. A general mathematical framework for representing soil organic matter dynamics. *Ecological Monographs*, 85(4):505–524, Nov. 2015. doi: 10.1890/15-0361.1. URL https://doi.org/10.1890/15-0361.1.

C. A. Sierra, V. Ceballos-Núñez, H. Metzler, and M. Müller. Representing and understanding the carbon cycle using the theory of compartmental dynamical systems. *Journal of Advances in Modeling Earth Systems*, 10(8):1729–1734, Aug. 2018. doi: 10.1029/2018ms001360. URL https://doi.org/10.1029/2018ms001360.

M. S. Smith, W. W. Frye, and J. J. Varco. Legume winter cover crops. In *Advances in Soil Science*, pages 95–139. Springer New York, 1987. doi: 10.1007/978-1-4612-4790-6_3. URL https://doi.org/10.1007/978-1-4612-4790-6_3.

J.-F. Soussana, S. Lutfalla, F. Ehrhardt, T. Rosenstock, C. Lamanna, P. Havlík, M. Richards, E. L. Wollenberg, J.-L. Chotte, E. Torquebiau, P. Ciais, P. Smith, and R. Lal. Matching policy and science: Rationale for the '4 per 1000 - soils for food security and climate' initiative. *Soil and Tillage Research*, 188:3–15, May 2019. doi: 10.1016/j.still.2017.12.002. URL https://doi.org/10.1016/j.still.2017.12.002.

R. F. Stallard. Terrestrial sedimentation and the carbon cycle: Coupling weathering and erosion to carbon burial. *Global Biogeochemical Cycles*, 12(2):231–257, June 1998. doi: 10.1029/98gb00741. URL https://doi.org/10.1029/98gb00741.

Statistical Office of the European Union. *Agriculture, forestry and fishery statistics: 2020 edition.* European Commission, 2020. doi: 10.2785/143455. URL https://data.europa.eu/doi/10.2785/143455.

R. D. Stewart, J. Jian, A. J. Gyawali, W. E. Thomason, B. D. Badgley, M. S. Reiter, and M. S. Strickland. What we talk about when we talk about soil health. *Agricultural and Environmental Letters*, 3(1):180033, Jan. 2018. doi: 10.2134/ael2018.06.0033. URL https://doi.org/10.2134/ael2018.06.0033.

G. Strang. *Introduction to linear algebra*. Cambridge Press, Wellesley, MA, 2016. ISBN 9780980232776.

J. P. M. Syvitski, C. J. Vörösmarty, A. J. Kettner, and P. Green. Impact of humans on the flux of terrestrial sediment to the global coastal ocean. *Science*, 308(5720):376–380, Apr. 2005. doi: 10.1126/science.1109454. URL https://doi.org/10.1126/science.1109454.

L. Tierney, A. J. Rossini, N. Li, and H. Sevcikova. *snow: Simple Network of Workstations*, 2021. URL https://CRAN.R-project.org/package=snow. R package version 0.4-4.

United Nations. *United Nations Framework Convention on Climate Change*. FCCC/INFORMAL/84. GE.05-62220 (E) 200705, 1992.

United Nations Environment Programme. *21 issues for the 21st century : results of the UNEP foresight process on emerging environmental issues*. United Nations Environment Programme, Nairobi, 2012. ISBN 9789280731910.

United Nations Environment Programme. Gemstat database of the global environment monitoring system for freshwater (gems/water) programme, 2018a. URL https://gemstat.org/data/. Accessed: 2021-11-15.

United Nations Environment Programme. Gemstat database of the global environment monitoring system for freshwater (gems/water) programme, 2018b. URL https://gemstat.org/data/. Accessed: 2021-11-15.

U.S.Government Acocuntability Office. Agricultural conservation: Usda's environmental quality incentives program could be improved to optimize benefits, 2017. URL https://www.gao.gov/assets/gao-17-225.pdf.

C. Utazi, J. Thorley, V. Alegana, M. Ferrari, K. Nilsen, S. Takahashi, C. Metcalf, J. Lessler, and A. Tatem. A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. *Statistical Methods in Medical Research*, 28(10-11):3226–3241, Sept. 2018. doi: 10.1177/0962280218797362. URL https://doi.org/10.1177/0962280218797362.

K. van Oost, G. Govers, T. A. Quine, G. Heckrath, J. E. Olesen, S. D. Gryze, and R. Merckx. Landscape-scale modeling of carbon cycling under the impact of soil redistribution: The role of tillage erosion. *Global Biogeochemical Cycles*, 19(4):n/a–n/a, Nov. 2005. doi: 10.1029/2005gb002471. URL https://doi.org/10.1029/2005gb002471.

K. van Oost, T. A. Quine, G. Govers, S. D. Gryze, J. Six, J. W. Harden, J. C. Ritchie, G. W. McCarty, G. Heckrath, C. Kosmas, J. V. Giraldez, J. R. M. da Silva, and R. Merckx. The impact of agricultural soil erosion on the global carbon cycle. *Science*, 318(5850):626–629, Oct. 2007. doi: 10.1126/science.1145724. URL https://doi.org/10.1126/science.1145724.

N.-C. Vavlas, T. W. Waine, J. Meersmans, P. J. Burgess, G. Fontanelli, and G. M. Richter. Deriving wheat crop productivity indicators using sentinel-1 time series. *Remote Sensing*, 12(15):2385, July 2020. doi: 10.3390/rs12152385. URL https://doi.org/10.3390/rs12152385.

A. Veloso, S. Mermoz, A. Bouvet, T. L. Toan, M. Planells, J.-F. Dejoux, and E. Ceschia. Understanding the temporal behavior of crops using sentinel-1 and sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*, 199:415–426, Sept. 2017. doi: 10.1016/j.rse.2017.07.015. URL https://doi.org/10.1016/j.rse.2017.07.015.

G. Verstraeten, K. Oost, A. Rompaey, J. Poesen, and G. Govers. Evaluating an integrated approach to catchment management to reduce soil loss and sediment pollution through modelling. *Soil Use and Management*, 18(4):386–394, Jan. 2006. doi: 10.1111/j.1475-2743.2002.tb00257.x. URL https://doi.org/10.1111/j.1475-2743.2002.tb00257.x.

M. Vreugdenhil, C. Navacchi, B. Bauer-Marschallinger, S. Hahn, S. Steele-Dunne, I. Pfeil, W. Dorigo, and W. Wagner. Sentinel-1 cross ratio and vegetation optical depth: A comparison over europe. *Remote Sensing*, 12(20):3404, Oct. 2020. doi: 10.3390/rs12203404. URL https://doi.org/10.3390/rs12203404.

D. Walling. Human impact on land–ocean sediment transfer by the world's rivers. *Geomorphology*, 79(3-4):192–216, Sept. 2006. doi: 10.1016/j.geomorph.2006.06.019. URL https://doi.org/10.1016/j.geomorph.2006.06.019.

J. E. Walsh, R. L. Thoman, U. S. Bhatt, P. A. Bieniek, B. Brettschneider, et al. The high latitude marine heat wave of 2016 and its impacts on alaska. *Bulletin of the American Meteorological*

Society, 99(1):S39–S43, Jan. 2018. doi: 10.1175/bams-d-17-0105.1. URL https://doi.org/10.1175/bams-d-17-0105.1.

H. Wang and M. G. Ranalli. Low-rank smoothing splines on complicated domains. *Biometrics*, 63(1):209–217, Nov. 2006. doi: 10.1111/j.1541-0420.2006.00674.x. URL https://doi.org/10.1111/j.1541-0420.2006.00674.x.

Z. Wang, S. Doetterl, M. Vanclooster, B. van Wesemael, and K. van Oost. Constraining a coupled erosion and soil organic carbon model using hillslope-scale patterns of carbon stocks and pool composition. *Journal of Geophysical Research: Biogeosciences*, 120(3):452–465, Mar. 2015. doi: 10.1002/2014jg002768. URL https://doi.org/10.1002/2014jg002768.

Z. Wang, T. Hoffmann, J. Six, J. O. Kaplan, G. Govers, S. Doetterl, and K. van Oost. Human-induced erosion has offset one-third of carbon emissions from land cover change. *Nature Climate Change*, 7(5):345–349, Apr. 2017. doi: 10.1038/nclimate3263. URL https://doi.org/10.1038/nclimate3263.

W. M. Washington, L. Buja, and A. Craig. The computational future for climate and earth system models: on the path to petaflop and beyond. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1890):833–846, Dec. 2008. doi: 10.1098/rsta.2008.0219. URL https://doi.org/10.1098/rsta.2008.0219.

D. J. Weiss, T. C. D. Lucas, M. Nguyen, A. K. Nandi, D. Bisanzio, K. E. Battle, E. Cameron, K. A. Twohig, D. A. Pfeffer, J. A. Rozier, H. S. Gibson, P. C. Rao, D. Casey, A. Bertozzi-Villa, E. L. Collins, U. Dalrymple, N. Gray, J. R. Harris, R. E. Howes, S. Y. Kang, S. H. Keddie, D. May, S. Rumisha, M. P. Thorn, R. Barber, N. Fullman, C. K. Huynh, X. Kulikoff, M. J. Kutz, A. D. Lopez, A. H. Mokdad, M. Naghavi, G. Nguyen, K. A. Shackelford, T. Vos, H. Wang, D. L. Smith, S. S. Lim, C. J. L. Murray, S. Bhatt, S. I. Hay, and P. W. Gething. Mapping the global prevalence, incidence, and mortality of plasmodium falciparum, 2000–17: a spatial and temporal modelling study. *The Lancet*, 394(10195):322–331, July 2019. doi: 10.1016/s0140-6736(19)31097-9. URL https://doi.org/10.1016/s0140-6736(19)31097-9.

W. Wischmeier, D. Smith, U. S. Science, E. Administration, and P. U. A. E. Station. *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning*. Agriculture handbook. Department of Agriculture, Science and Education Administration, 1978.

W. H. Wischmeier. A rainfall erosion index for a universal soil-loss equation. *Soil Science Society of America Journal*, 23(3):246–249, May 1959. doi: 10.2136/sssaj1959.03615995002300030027x. URL https://doi.org/10.2136/sssaj1959.03615995002300030027x.

R. Wolfinger. Laplace's approximation for nonlinear mixed models. *Biometrika*, 80(4):791–795, 1993. doi: 10.1093/biomet/80.4.791. URL https://doi.org/10.1093/biomet/80.4.791.

S. Wood. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edition, 2017.

S. N. Wood. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, 62(4):1025–1036, May 2006. doi: 10.1111/j.1541-0420.2006.00574.x. URL https://doi.org/10.1111/j.1541-0420.2006.00574.x.

S. N. Wood. P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. *Statistics and Computing*, 27(4):985–989, May 2016. doi: 10.1007/s11222-016-9666-x. URL https://doi.org/10.1007/s11222-016-9666-x.

J. Y. Xia, Y. Q. Luo, Y.-P. Wang, E. S. Weng, and O. Hararuk. A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, 5(5):1259–1271, 2012. doi: 10.5194/gmd-5-1259-2012. URL https://gmd.copernicus.org/articles/5/1259/2012/.

C. Zhou, K. Wang, and D. Qi. Attribution of the july 2016 extreme precipitation event over china's wuhang. *Bulletin of the American Meteorological Society*, 99(1):S107–S112, Jan. 2018. doi: 10.1175/bams-d-17-0090.1. URL https://doi.org/10.1175/bams-d-17-0090.1.

Résumé long en français

Dans le contexte du système climatique terrestre, les sols constituent une ressource naturelle essentielle. Bon nombre des services écosystémiques fournis par les sols sont essentiels au maintien des moyens de subsistance de l'homme et sont affectés par le changement climatique. Il s'agit notamment de la production de nourriture, d'eau douce, d'énergie et d'habitats pour la biodiversité. La demande humaine pour ces services, et par conséquent la pression sur les sols, tend à augmenter avec la croissance démographique mondiale prévue (IPCC, 2019). En outre, les sols de la planète contiennent entre 1 500 et 2 400 pétagrammes de carbone (PgC), soit plus que l'atmosphère (589 PgC) et l'océan de surface (900 PgC) réunis (Ciais et al., 2013). Par conséquent, même de petites perturbations dans les réservoirs de carbone du sol peuvent avoir un impact substantiel sur leurs émissions vers l'atmosphère. Par exemple, des mesures de gestion plus durables, telles que la conservation des résidus dans les systèmes de culture, peuvent augmenter la teneur en carbone organique du sol et contribuer à séquestrer le carbone de l'atmosphère (Robert, 2001), et atténuer une menace majeure pour les sols, à savoir l'érosion des sols.

L'érosion accélérée des sols est le processus le plus répandu de dégradation des sols causée par les activités agricoles (Robert, 2005), aux côtés du compactage, l'épuisement des nutriments, la décomposition des agrégats, et d'autres phénomènes. Selon Hillel (2004), le processus physique de l'érosion des sols peut être décrit en trois étapes : i) Tout d'abord, les particules sont détachées du sol. Les particules les plus vulnérables sont celles de la surface du sol, qui est la partie la plus fertile du profil en raison des grandes quantités de matière organique (humus) et de micro-organismes qu'y sont présents. La perte de fertilité causée par le détachement des particules de la couche arable nécessite généralement l'utilisation d'engrais chimiques, ce qui entraîne une menace croissante de pollution des eaux souterraines. ii) Deuxièmement, les particules détachées sont transportées à travers le paysage ; et iii) Troisièmement, les particules sont déposées à un endroit différent de celui où elles se trouvaient à l'origine. Par ailleurs, les auteurs incluent parfois la décomposition des macroagrégats comme un processus intermédiaire entre le premier et le deuxième processus précédents (Lal, 2005). L'érosion du sol peut être déclenchée par l'eau ou le vent, le premier étant le plus courant.

Le flux de particules à travers le paysage est guidé par la gravité et suit les caractéristiques géomorphologiques du terrain. Par conséquent, l'érosion du sol peut avoir des impacts sur site, tels que la perte de productivité, et des impacts hors site, tels que l'eutrophisation des rivières et des réservoirs. C'est pourquoi l'érosion des sols est considérée comme un mécanisme de transport essentiel permettant d'acheminer les effets environnementaux des produits chimiques agricoles vers des zones plus éloignées (Hillel, 2004). Chaque année, 20 à 30 Gt de sols sont perdus dans le monde à cause de l'érosion hydrique, les taux par unité de surface variant en fonction de la zone climatique. Dans les régions tempérées, les valeurs moyennes sont généralement de 10 t ha^{-1} an^{-1}, mais peuvent atteindre 20 t ha^{-1}

an^{-1} dans les zones de cultures vallonnées. Dans les régions tropicales, qui concentrent les taux d'érosion les plus élevés au monde, les valeurs varient souvent entre 10 et 20 t ha^{-1} an^{-1}, et peuvent atteindre 50 t ha^{-1} an^{-1} dans les cas les plus extrêmes (FAO, 2014; van Oost et al., 2007). Ces valeurs sont corroborées par les résultats plus récents de Borrelli et al. (2018a), estimant une perte moyenne de sol de 35 Gt en 2001 et de 35,9 Gt en 2012, les taux spécifiques les plus élevés étant enregistrés en Amérique du Sud, en Afrique et en Asie, respectivement.

La convergence des prévisions d'érosion dans la littérature implique que, quelle que soit la base de données utilisée, les valeurs sont beaucoup plus élevées que les taux de formation des sols (Julien, 2010). Ces derniers sont généralement inférieurs à 1 t ha^{-1} an^{-1}, avec une médiane de 0,15 t ha^{-1} an^{-1}. Un tel déséquilibre entre la formation et la perte de sol signifie que les sols du monde pourraient être traités comme une ressource naturelle non renouvelable plutôt que renouvelable. En outre, à long terme, pendant des décennies ou des siècles, ces taux d'érosion sont susceptibles de directement réduire le rendement des cultures en raison de la diminution de la capacité de rétention en eau et de la réduction de l'espace réservé aux racines (FAO, 2014).

Il existe plusieurs liens entre l'érosion des sols et le cycle du carbone (Lal, 2005). Sur site, il est admis que tous les mécanismes conduisent à une augmentation des émissions atmosphériques. Ces mécanisms incluent notamment l'élimination de l'argile et du carbone organique du sol, l'augmentation des taux de minéralisation en raison des modifications de l'humidité du sol, l'exposition du sous-sol et la décomposition des agrégats. Hors site, certains effets, tels que le transport vers d'autres éléments du paysage, comme les plaines d'inondation et les systèmes aquatiques, pourraient conduire à la protection et à la séquestration du carbone. L'intensité de chacun de ces effets a conduit à un long débat scientifique pour comprendre si l'érosion des sols correspond à une source ou à un puits de carbone atmosphérique, avec des résultats allant dans des directions différentes. La modélisation de la relation entre l'érosion des sols et le cycle du carbone organique est toutefois une tâche complexe.

Ces dernières années, la popularisation d'ordinateurs puissants a permis une large diffusion de différentes approches de modélisation. L'examen systématique de Borrelli et al. (2021) a montré que 1 697 articles ont publié 3 030 applications de modélisation de l'érosion entre 1994 et 2017. Leur analyse a montré la prédominance des études sur l'érosion hydrique à l'échelle nationale et une tendance temporelle claire à l'augmentation des applications, allant de 55 en 1994-1997 à 340 en 2014-2017. Les auteurs ont également trouvé 435 modèles d'érosion différents, avec une nette prédominance des approches basées sur l'USLE, correspondant à 41% de toutes les applications. L'une des raisons de la grande popularité des méthodes basées sur l'USLE est un compromis relativement bon entre les exigences en matière de données d'entrée et la fiabilité des estimations de la perte de sol. Cependant, certaines limitations existent : i) les applications constituent souvent des extrapolations de l'échelle originale locale à laquelle le modèle a été développé, et ii) le manque de représentation du transport et du dépôt de sédiments (Quine and van Oost, 2020). La deuxième limitation est souvent résolue en incluant un schéma externe d'acheminement des sédiments dans le cadre de modélisation, par exemple (Rompaey et al., 2001).

Historiquement, le développement des modèles d'érosion s'est fait parallèlement à celui des modèles de surface terrestre. Les modèles de surface terrestre visent à intégrer de multiples équations pour simuler les processus terrestres, notamment : les processus physiques et les transformations chimiques (par exemple, la croissance des feuilles), les interactions entre les colonnes (par exemple, la chute des précipitations), l'échange entre les composants (par exemple, l'eau précipitée qui pénètre dans le sol), l'échange de chaleur et le mouvement des substances (par exemple, l'écoulement de l'eau). Tout au long de ces processus, les calculs sont limités par des lois physiques fondamentales, telles que la conservation de la masse et de l'énergie (Gettelman and Rood, 2016). Selon Doetterl et al. (2016), aucun modèle de surface terrestre ne contient actuellement une représentation suffisamment détaillée de la relation entre le carbone organique du sol et l'érosion, le transport et le dépôt. Les auteurs affirment que cela est dû à des données d'entrée inadéquates, à des problèmes de généralisation des paramètres du modèle, à une compréhension insuffisante des processus d'érosion et à d'importantes exigences en matière de calcul.

Plusieurs travaux dans la littérature ont visé à coupler l'érosion, le transport et le dépôt au cycle biogéochimique du carbone, avec des approches qui varient en termes de détail, de complexité et d'échelle. Lal (2003), par exemple, a supposé un acheminement fixe des sédiments pour estimer les valeurs globales agrégées du déplacement du carbone. Cette approche est similaire à celle de Chappell et al. (2015), qui a pris en compte l'érosion comme un flux de carbone supplémentaire, mais n'a pas tenu compte du transport de particules. Dans l'approche spatialement explicite de Lugato et al. (2016), l'érosion a également été prise en compte comme un flux supplémentaire, et le transport de particules vers les rivières a été supposé être une part constante du sol érodé total dans chaque cellule de la grille. Wang et al. (2017) a également adopté un modèle de transport spatialement explicite, dans lequel le paysage a été divisé en régions appelées "bassins versants virtuels" pour tenir compte d'un flux de transport supplémentaire. À l'échelle continentale, Borrelli et al. (2018a) a couplé les données sur le carbone organique du sol à un modèle d'érosion représentant le transport de particules d'une manière distribuée basée sur les processus. Cette approche est toutefois agrégée et n'inclut aucun autre élément du cycle du carbone. À l'échelle du bassin versant, Nadeu et al. (2015) a inséré un schéma de routage spatialement distribué dans un modèle de carbone détaillé dans le centre de la Belgique. Toutefois, cette approche est limitée par l'échelle à laquelle le modèle peut être appliqué.

Pour résoudre les problèmes mentionnés ci-dessus, Naipal et al. (2020) a proposé le modèle CE-DYNAM pour la partie non alpine du bassin du Rhin, une région dont la superficie est égale à 185 000 km2. Ce modèle combine des modules d'érosion et de transport à un émulateur de la dynamique du carbone du sol de n'importe quel modèle de surface terrestre basé sur CENTURY (Parton et al., 1983). CE-DYNAM utilise une approche basée sur le modèle RUSLE adaptée à la modélisation de l'érosion à grande échelle et à une résolution spatiale grossière (Naipal et al., 2015), et dont le module de transport est un schéma de routage qui suit la topographie pour redistribuer les particules de sol (Naipal et al., 2016). Ces éléments sont formellement incorporés dans la dynamique du modèle en tant que flux supplémentaires entre les bassins au-delà de ceux initialement présents dans la cinétique de premier ordre de CENTURY.

Dans sa première version, CE-DYNAM était adapté à une utilisation à l'échelle du bassin versant et ne pouvait pas être correctement adapté à des domaines plus vastes pour plusieurs raisons. Tout d'abord, les temps d'exécution allaient de quelques heures à quelques jours à l'échelle du bassin versant, ce qui indique que les calculs pour des domaines plus vastes pourraient prendre des semaines. Une telle limitation empêchait également l'exécution d'une procédure d'étalonnage puisque les routines d'optimisation nécessitent souvent plusieurs évaluations de fonctions. Deuxièmement, la disponibilité de la mémoire de l'ordinateur était un facteur limitant puisque tous les ensembles de données pour l'ensemble du domaine devaient être chargés et lus avant le traitement. Troisièmement, la mise en œuvre du code de CE-DYNAM contenait des détails qui le rendaient impraticable pour une application lorsque le nombre de sous-bassins versants dans le domaine augmentait. Enfin, l'intensité des flux latéraux devait être imposée et non calibrée avec des données observées, ce qui pouvait générer des résultats irréalistes. Tous ces problèmes étaient liés et indiquaient que, malgré son potentiel, des travaux supplémentaires étaient nécessaires pour améliorer CE-DYNAM.

Dans ce contexte, la présente thèse de doctorat avait plusieurs objectifs. Le premier objectif (**Chapitre 2**) était de proposer une formulation alternative pour CE-DYNAM afin de permettre la mise à l'échelle de ses calculs au niveau continental. Ceci a été fait en formalisant la formulation du modèle, en représentant la dynamique du premier ordre du modèle sous forme de matrice, et en introduisant une procédure de calibration basée sur des mesures de décharge de sédiments. Dans ce chapitre, une deuxième version de CE-DYNAM (i.e. v2) a été présentée, qui consiste en sa version initiale (v1) plus la forme matricielle et quelques modifications. Le premier chapitre traite de *modélisation mécaniste*.

Le deuxième objectif (**Chapitre 3**) était de générer la première carte par satellite des cultures de couverture en Europe. L'adoption de cultures de couverture étant une mesure de gestion ayant un impact direct sur le cycle du carbone organique du sol et sur la prévention de l'érosion, ce chapitre visait à mieux comprendre leur répartition spatiale à l'échelle continentale. Le deuxième chapitre traite de *la modélisation empirique*.

Enfin, le troisième objectif (**Chapitre 4**) était d'évaluer les impacts potentiels des cultures de couverture sur les stocks de carbone du sol et l'exportation de carbone organique particulaire vers les océans en Europe en utilisant CE-DYNAM. Ce travail a combiné les développements des chapitres 1 et 2, et une troisième version de CE-DYNAM (i.e., v3) a été présentée. Cette version se compose de la v2 et de plusieurs modifications majeures ajoutées pour améliorer la représentation physique des processus d'érosion, de transport et de dépôt. Le troisième chapitre traite de *l'évaluation des alternatives politiques*.