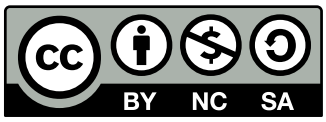
Aalto University
School of Science

Master's Programme in Mathematics and Operations Research

# Bayesian modeling of a carbon farming experiment

**Velma Pohjonen**

Master's Thesis
2025



Aalto University
School of Science

**Author** Velma Pohjonen

**Title** Bayesian modeling of a carbon farming experiment

**Degree programme** Mathematics and Operations Research

**Major** Systems and Operations Research

**Supervisor** Prof. Pauliina Ilmonen

**Advisor** Dr. Julius Vira

**Collaborative partner** Finnish Meteorological Institute

**Date** 13 September 2025 **Number of pages** 52 **Language** English

**Abstract**

Soil is one of the largest carbon reservoirs in the world. However, the conversion of land into agricultural fields has led to a continuous loss of soil carbon stocks. The release of carbon from croplands not only increases greenhouse gas emissions but can also lead to poorer harvests. To preserve soil carbon stocks and secure food production, carbon farming practices that reduce the loss of carbon stocks or even increase the sequestration of carbon in the soil have been developed. However, assessing the impact of carbon farming methods is difficult, as there is a great deal of uncertainty associated with measuring soil carbon stocks.

This thesis focuses on modeling the soil carbon stocks of fields participating in a carbon farming experiment and assessing the impact of carbon farming practices on the fields. The fields are divided into treatment fields that are cultivated using carbon farming practices and control fields that are cultivated using more conventional methods. Soil carbon stocks are first modeled at individual time points, with an assumption that carbon content decreases exponentially with depth. Using Bayesian methods, the model is conditioned to measurements of carbon concentration on the surface of the soil. Carbon stocks are then simulated for the years 1900-2025, using Yasso, a soil carbon model. With Bayesian methods, the simulation is conditioned on the modeled carbon stocks, as well as estimates of plant biomass and other carbon inputs in the fields.

The study then examines the carbon input into the soil, photosynthesis, and changes in carbon stocks during the carbon farming experiment. The fields were grouped according to their carbon farming practice in fields that used nutrient-rich organic amendments, cover crops, and other methods. On average, organic amendments increased the carbon input to the fields by 25% and reduced the loss of carbon stocks by 55%, compared to the carbon input to the control fields. No significant differences were observed in the photosynthesis rates between the treatment and control fields, with any carbon farming practices. Overall, carbon stocks of the treatment fields decreased by 0.48% annually during the carbon farming experiment, while the stocks of control fields decreased by 0.54%. However, uncertainties were large throughout the results, which weakens their statistical significance.

**Keywords** Bayesian modeling, carbon farming, soil organic carbon, Yasso

Aalto-yliopisto
Perustieteiden
korkeakoulu

**Tekijä** Velma Pohjonen

**Työn nimi** Hiiliviljelykokeen bayesiläinen mallinnus

**Koulutusohjelma** Matematiikka ja operaatiotutkimus

**Pääaine** Systeemi- ja operaatiotutkimus

**Työn valvoja** Prof. Pauliina Ilmonen

**Työn ohjaaja** FT Julius Vira

**Yhteistyötaho** Ilmatieteen laitos

**Päivämäärä** 13.9.2025 **Sivumäärä** 52 **Kieli** englanti

**Tiivistelmä**

Maaperä on yksi maailman suurimmista hiilivarastoista. Maa-alueiden ottaminen viljelykäyttöön on kuitenkin johtanut siihen, että maaperä menettää hiiltä jatkuvasti. Hiilen vapautuminen viljelymailta paitsi lisää kasvihuonepäästöjä, voi myös johtaa huonompiin satoihin. Maaperän hiilivaraston säilyttämiseksi ja ruoantuotannon turvaamiseksi on kehitetty hiiliviljelymenetelmiä, joiden on tarkoitus vähentää hiilen vapautumista, tai jopa edistää sen sitoutumista maaperään. Hiiliviljelymenetelmien vaikutusten arviointi on kuitenkin vaikeaa, sillä maaperän hiilivarastojen mittaamiseen liittyy paljon epävarmuutta.

Tässä työssä mallinnetaan hiiliviljelytutkimukseen osallistuneiden peltojen maahiilivarastoja, ja sitä kautta tutkitaan hiiliviljelymenetelmien vaikutusta peltoihin. Pellot on jaettu hiiliviljelymenetelmin viljeltyihin käsittelypeltoihin ja tavanomaisin menetelmin viljeltyihin kontrollipeltoihin. Hiilivarastot mallinnetaan ensin yksittäisille ajankohdille niin, että hiilipitoisuuden oletetaan vähenevän syvyyden suhteen eksponentiaalisesti. Malli ehdollistetaan bayesialaisten menetelmien avulla peltojen pintamaasta otetuille hiilipitoisuusmittauksille. Hiilivarastoja simuloidaan sen jälkeen Yasso-maahiilimallin avulla vuosille 1900-2025. Simulointi ehdollistetaan bayesialaisia menetelmiä käyttäen mallinnetuille maahiilivarastoille, sekä arvioille pelloilla tapahtuneesta yhteytyksestä ja muusta hiilisyötteestä.

Työssä tarkastellaan hiiliviljelykokeen aikana tapahtunutta hiilisyötettä maaperään, yhteytystä sekä hiilivarastojen muutosta. Pellot on ryhmitelty hiiliviljelymenetelmien mukaan ravinnerikkaita orgaanisia lannoitteita, peitekasveja ja muita menetelmiä käyttäneisiin peltoihin. Orgaanisten lannoitteiden havaittiin lisänneen peltojen hiilisyötettä keskimäärin 25% kontrollipeltojen hiilisyötteeseen verrattuna. Niiden käyttö myös vähensi hiilivaraston laskua keskimäärin 55%. Yhteytyksessä ei nähty merkittäviä eroja käsittely- ja kontrollipeltojen välillä minkään hiiliviljelymenetelmän osalta. Kokonaisuudessaan käsittelypeltojen hiilivarastot laskivat hiiliviljelykokeen aikana vuosittain 0.48% ja kontrollipeltojen 0.54%. Tuloksissa oli kuitenkin suuria epävarmuuksia, mikä heikentää niiden tilastollista merkitsevyyttä.

**Avainsanat** Bayesialainen mallinnus, hiiliviljely, maaperän orgaaninen hiili, Yasso

# Preface

First and foremost, I want to thank my instructor Julius for all the help and support he gave me during my thesis work. Thank you for giving me such an interesting topic to work with. Thank you for all the patience, for the long coding sessions, and for one more time answering my questions about $\gamma$. Working on this thesis with you has truly been a joy, and I couldn't have done it without you.

Thank you, Pauliina, for the great supervision. Thank you for the countless meetings, for the overwhelmingly positive feedback after each text snippet you saw, and for having so much faith in me. After each meeting we had, I had more confidence that one day this thesis was going to be completed.

To our group leader Liisa, thank you, first of all, for giving me a job, and for reminding me that the best type of a thesis is a ready thesis. Thank you for leading our group with such warmth. To the whole community of FMI, thank you for making our workplace a place I am happy to go to. For filling my days with laughter and ridiculous conversations, and for tolerating my unreasonably strong opinions about literally anything.

During my studies, many communities have welcomed me with open arms. In the Guild of Physics I was lucky to meet wonderful people and leave my mark on the Guild's sports, wellness, and senior activities. Working in the board of AYY sparked my interest in social advocacy. In Fyysikkospeksi I got to fulfill my dreams as a dancer. In BY, I truly felt like home and I left with countless stories that I can use to shock people. Thank you for making my studies a time I will remember fondly.

To all my friends, thank you for occasionally distracting me from my thesis. Thank you for the summer trips, the wine nights, the barn dances and the singing sessions. Thank you for making unforgettable memories with me, whether they are a Sanna Marin themed party or just eating strawberries on the beach.

And finally, to my family. Thank you for all your support. For making me the person I am, and for loving that silly girl.

Helsinki, December 8, 2025

Velma Pohjonen

# Contents

# Symbols and abbreviations

## Symbols

| | |
|---|---|
| $U(a, b)$ | Uniform distribution with minimum $a$ and maximum $b$ |
| $N(\mu, \sigma^2)$ | Normal distribution with mean $\mu$ and standard deviation $\sigma$ |
| $N^+(\sigma^2)$ | Half-normal distribution with standard deviation $\sigma$ |

## Abbreviations

| | |
|---|---|
| C | carbon |
| $CO_2$ | Carbon dioxide |
| GPP | gross primary production |
| kg C | kilogram of carbon |
| NPP | net primary production |
| SOC | soil organic carbon |
| SOM | soil organic matter |

# 1 Introduction

Soil is one of the largest carbon reservoirs on the planet. The amount of carbon stored in the top meter of soil is estimated to be around 1500 Gt, making soil the major terrestrial carbon stock (Batjes, 2014). The carbon reservoir in soils exceeds the combined carbon stocks of vegetation and the atmosphere (Paustian et al., 2019; Georgiou et al., 2022).

However, soils are losing the carbon stored in them. In 2019, 45% of the world's soils were in agricultural use (Paustian et al., 2019). Agricultural soils generally contain less carbon than soils in their natural state: In agriculture, the sequestered carbon in the crop is constantly harvested and taken away from the field, preventing it from being stored in the soil (Paustian et al., 2019). Total plant growth is often smaller in agricultural lands and annual crops have less below-ground biomass such as roots (Field, 2001; Crews and Rumsey, 2017). According to Davidson and Ackerman (1993), around 20-40% of soil carbon is lost when untilled soil is brought into cultivation.

The release of carbon from croplands not only increases greenhouse gas emissions, but can also lead to poorer harvests (Lal, 2004). In order to mitigate climate change and strengthen food security, an initiative to increase global soil carbon stocks by 0.4% annually was launched at the 2015 United Nations Climate Change Conference (Minasny et al., 2017). Agricultural methods known as carbon farming, which aim to increase carbon sequestration in the soil, could be one way of increasing the stocks. However, there are different views on the effectiveness of carbon farming: according to Paustian et al. (2019), the use of carbon farming methods could make soil an important carbon sink, while Günther et al. (2024) stated that models often overestimate the effect of carbon farming. The impact of carbon farming on soil carbon stocks should therefore be further investigated. However, there are challenges associated with assessing the size of soil carbon stocks. Soil samples used to determine the carbon content of soil are expensive, and the carbon content can vary a lot within the same agricultural field (Abdullahi et al., 2018; Goidts et al., 2009). More accurate modeling could give insights on the effects of carbon farming.

In this thesis, the impact of carbon farming is examined, and new methods for soil carbon modeling are presented. We model carbon stocks of mineral soils in farms that took part in a carbon farming experiment described in Mattila et al. (2022), and examine their response to carbon farming. The fields where carbon farming was practiced are compared with those without carbon farming. The comparison is done in terms of photosynthesis on the fields, carbon input into the soil, and changes in carbon stocks during the experiment. Different carbon farming methods are mostly treated as a single group, however, usage of cover crops and organic fertilizers are also examined separately.

Soil carbon modeling is done by combining modeling and measured data. Carbon stocks in the experiment fields are modeled based on initial assumptions, which are then updated based on measurements of carbon content. The updating is done using Bayesian methods. A subset of fields with more measurements are used to evaluate a general distribution of carbon in soil, to model carbon stocks of fields with less measurements. Long-term carbon stocks are then simulated with the soil carbon model

Yasso, and the simulation is guided by data on carbon inputs and carbon stocks of individual years.

# 2 Background

## 2.1 Carbon in soil

Soil organic matter (SOM) refers to the plant and microbial residues in soil that have decayed long enough to be unrecognizable (Chapin III et al., 2011). When the dead organic matter is still recognizable, it is called litter (Chapin III et al., 2011). Soil organic carbon (SOC) refers to the carbon that is in the compounds of SOM (Food and Agriculture Organization of the United Nations, 2025). Carbon enters the soil mainly through organic compounds released by roots, or through the decomposition of litter (European Union, 2011).

Decomposition is a process in which organic compounds or remains of organisms break down into new organic compounds, nutrients, and carbon dioxide ($CO_2$). The most important decomposers are fungi and bacteria. The rate of decomposition primarily depends on susceptibility of a substrate to decomposition (e.g. size of molecules and chemical bonds), characteristics of the decomposer community (e.g. microbial activity and biomass of microbes), and physical environment (e.g. temperature and moisture). (Chapin III et al., 2011)

SOM consists of countless compounds, including sugars, amino acids, cellulose, lignin, etc. The rate at which compounds break down varies, as do their chemical structures and sizes. The dark and hardly decaying type of SOM is called humus. It is formed slowly as a result of decomposition. Humus is not chemically uniform in structure, but is a very irregular substance consisting of various compounds. Humus can stay in the soil from ten years up to thousands of years. (Chapin III et al., 2011)

Most decomposition occurs near the soil surface, as this is where most organic residues end up. The soil is therefore richest in carbon near the surface: in mineral soil, the top 20 cm of soil contains around half of the soil's SOC. (Chapin III et al., 2011)

## 2.2 Carbon cycle of cultivated mineral soils

Most agricultural soils are mineral soils, which are fairly poor in carbon (Paustian et al., 2019). According to Wilson (2019), organic matter makes up less than 20% of the mineral soil mass, while in the research by Paustian et al. (2019), the same proportion is considered to be around 0-10%. Still, the carbon content (mass of carbon per mass of soil) of most mineral soils is less than 5%, even in the topsoil (Wilson, 2019). Similar to most croplands, the cultivated mineral soils are losing the carbon stored in them: between the years 1974 and 2009, the cropland mineral soils of Finland have, on average, lost 220 kg C / ha / year from the top 15 cm of the soil (Heikkinen et al., 2013). This accounts for around 0.4% yearly decrease in the carbon stocks of the topsoil (Heikkinen et al., 2013).

Changes in the carbon stock of agricultural soil depend mainly on how much carbon enters the soil in the form of plant residues and organic amendments, and how much is lost through decomposition (Paustian et al., 2017). In croplands, around 40-60% of the above-ground biomass of most modern crops is harvested, therefore reducing

the amount of plant residues available for the soil (Hay, 1995). This is considered the main reason for the decrease of SOM in agricultural soils (Leifeld, 2013).

## 2.3 Carbon farming

Carbon farming refers to agricultural management practices that intend to increase carbon sequestration into the soil, or decrease carbon emissions from the soil (Paustian et al., 2019). The goal of carbon farming is to decrease the amount of atmospheric carbon, but it can also have other positive effects, such as improved soil health and soil fertility (Paustian et al., 2019). There are various carbon farming methods, such as usage of cover crops, addition of organic amendments, and adding leys in crop rotation (Mattila et al., 2022). Carbon farming methods have several ways of working, for instance, lengthening photosynthesis period or increasing carbon inputs into the soil. This thesis pays special attention to two carbon farming practices: cover crops and organic soil amendments.

### 2.3.1 Cover crops

Sowing cover crops is a carbon farming strategy that has traditionally been used, for example, to prevent erosion and leaching of nutrients, and to improve nitrogen sequestration into the soil (Kaye and Quemada, 2017). With this method, another crop is sown on the field, in addition to the main crop. There are many species that can be used as cover crops, such as legumes, grasses and cereals, as well as mixtures of different species (European Commission, 2022; Mirsky et al., 2009). The cover crop can be sown after the main crop is harvested, or it can grow between the rows of the main crop or under it (Natural Resources Institute Finland, 2025; European Commission, 2022). Using cover crops lengthens the photosynthesis period, which can increase the primary production of a cropland (Mattila et al., 2022; Chenu et al., 2019).

### 2.3.2 Organic soil amendments

Organic matter additions is another example of a carbon farming technique. This method aims to increase SOC content by adding carbon into the soil in the form of organic materials, such as manure or compost. This can also improve soil properties such as nutrient availability, which can lead to increased vegetation, and therefore, increased photosynthesis (Paustian et al., 2019). In the carbon farming experiment under investigation, two types of organic amendments are considered: nutrient-rich (N-rich) and nutrient-poor amendments (Mattila et al., 2022). In this thesis, nutrient-rich amendments were chosen to represent organic soil amendments, as they were a more commonly used practice than the nutrient-poor amendments.

### 2.3.3 Other practices

In Mattila et al. (2022), there are also other carbon farming practices used, that in this thesis are not paid as close attention to as cover crops and nutrient-rich additions, but are mentioned here. In subsoiling, the soil is loosened in order to remove physical barriers blocking the root growth (Mattila, 2024). Ley in rotation means that every few years a field is cultivated with leys instead of cereals, as there is more below-ground biomass in grasslands than in arable lands (Nyameasem et al., 2024). Multi-variety ley refers to a strategy where a mixture of leys is sown into the field, as it can increase the rooting depth (Gantlett et al., 2024). Adaptive grazing imitates the natural grazing of herds, which has shown to enhance natural functions of soil and thus the flow of carbon into the soil (Johnson et al., 2022). Reduced tillage means decreasing the soil disturbance to improve soil health and reduce carbon leakage (Derpsch et al., 2010).

# 3 Data and the soil carbon model

## 3.1 Carbon Action experiment

In the Carbon Action experiment, described in Mattila et al. (2022), 105 farmers from all around the arable parts of Finland volunteered to implement carbon farming methods in their fields. The farmers received training on carbon farming methods, and after the training, chose a carbon farming practice to implement in their field. The farmers designed a 5-year carbon farming plan for the field, with the intention of increasing the carbon storage of the soil. Along these fields, called treatment sites, the farmers had control sites, where the carbon farming method was not implemented. A farming plan was made for the control site as well. Following these plans, the farmers implemented their chosen carbon farming practices in the treatment site, and the more conventional farming practices on the control site. As part of the experiment, soil samples were taken from both sites. Some farmers dropped out of the experiment, but 73 farmers remained until the end. (Mattila et al., 2022)

In order to be able to observe the effect of the carbon farming practice in question, the farming practices on the treatment and control sites differed only in terms of the carbon farming practice (Mattila et al., 2022). The cover crop practice was implemented by sowing a cover crop underneath an annual main crop (Susi et al., 2025). In the control site, the same main crop was cultivated, but no cover crop was sown underneath (Susi et al., 2025). With nutrient-rich amendments, the farmers added organic, nutrient-rich material into the treatment sites, and no amendments were added on the control sites (Mattila et al., 2022). If organic amendments were added with other practices than N-rich or N-poor amendments, they were added in both treatment and control sites (Mattila et al., 2022).

## 3.2 Data from Carbon Action

In this thesis, data collected as part of the Carbon Action experiment is used. This includes the farming measures taken in the fields: During the experiment, the farmers recorded which crop species were sown on the sites and when. The yields (mass of crops harvested from fields) were also recorded. Similarly, data of the mass of organic additives (e.g. manure) brought to the fields was collected.

In addition, data that was measured using soil samples is used. Soil samples were taken from both treatment and control sites. The intention was to collect the samples before and after the experiment, however, there is a lot of variation between the sampling dates: the first sampling was done during the years 2018-2020, and the second one during 2023-2024.

From some farms, the farmers were especially motivated and more data was collected: these are the so-called advanced farms. From advanced farms, soil samples were taken at several depths at 10 cm intervals, with the first one near the surface. The number of samples depends on the site: from some sites, samples were taken only to the depth of 40 cm, from others to a depth of up to 110 cm. From most advanced sites, replicates of the measurements were taken, meaning that the soil sampling was

repeated on the same site, so there are several sets of measurements from the same site. The carbon concentration of the soil at a given depth was determined using these samples. The masses of the 10 cm soil layers were also determined.

From most of the farms, less data was collected. These are called basic farms. From the basic sites, soil samples were taken only down to a 20 cm depth, and used to determine the carbon concentration of the soil. There were no replicates. The mass of the soil was not determined.

As this thesis focuses on mineral soils, only data from sites with a low-enough carbon concentration was used: If the carbon concentration was more than 10 % in any of the layers and in either year, the sites were considered organic soils and discarded. Similarly, if data were not available for both control and treatment sites, they were not included in the data.

In the end, there were 18 advanced farms and 64 farms in total whose data was used in this thesis. Out of the advanced farms, 5 used cover crops as their carbon farming method and 4 used nutrient-rich additions. Out of all the farms, 20 used cover crops and 6 nutrient-rich additions as their carbon farming method. The rest of the farms used the methods described in Sections 2.3.2 and 2.3.3, or a combination of different practices.

## 3.3 Data from Finnish Meteorological Institute

In addition to data from the carbon farming experiment, data from Finnish Meteorological Institute (FMI) is used. This includes weather data from the locations of the experimental fields. The data is from the years 1961-2025, which covers the available measurement history of FMI. For the years prior to 1961, average climate conditions are assumed, determined by the climatic conditions of years 1961-1990. Only data of temperature and precipitation are used.

In addition, data of photosynthesis rates on the fields are used. They are estimated using data of photosynthetically active radiation and vegetation index measured with Sentinel-2 satellites, which are fed into an empirical photosynthesis model (Vira et al., 2025). Net primary production (NPP), the fraction of carbon stored in the plants and not used in cellular respiration, is assumed to be 50% of the total photosynthetically fixed carbon (gross primary product, GPP).

## 3.4 Soil carbon model Yasso

Yasso is a model that describes the carbon cycle of soil: the stock of SOC and its decomposition (Viskari et al., 2022). Yasso was first published by Liski et al. (2005) as a carbon and decomposition model for forest soils. The second version of the model, Yasso07 (Tuomi et al., 2009), was tested by Karhu et al. (2011) for deforested sites that had been taken for agricultural use. Yasso07 seemed to predict changes in SOC quite well for these sites as well, and Yasso has since been used for agricultural sites as well. The newest version, Yasso20, was calibrated with multiple datasets across the world, with the restriction that the carbon stock of the soils were less than 20 kg C / $m^2$ (Viskari et al., 2022). In this thesis, Yasso20 is used, referred to below as Yasso.

Yasso models the carbon stock of soil over time. The stock is divided into labile and stable pools. The labile SOC is categorized into four compartments based on their solubility: compounds soluble in a polar solvent and those soluble in a non polar solvent; compounds hydrolysable in acid, and non-soluble and non-hydrolysable residues. The stabile SOC is represented by a fifth compartment, humus. Yasso is often run with one year as a timestep, but other steps are also possible. At each time step, Yasso simulates the decomposition of SOC from one compartment into another, or into $CO_2$. Decomposition rate is controlled by four factors: temperature, precipitation, litter composition, and the average diameter of woody litter. However, woody litter is only relevant when modeling forests. (Viskari et al., 2022)

As an input, Yasso is given weather data, which is used to define temperature and precipitation. Temperature is considered on a monthly resolution, precipitation as an annual average. Litter composition and average diameter of woody litter can also be given if they are measured. If not, default values can be used. In this thesis, we use values from Palosuo et al. (2015).

Yasso also needs an estimate of the initial carbon stock of the soil. Such data is often not available, especially if the purpose is to simulate carbon stock from the distant past. In this thesis, an initial estimate of the stock is given for Yasso, which is later updated using Bayesian methods and measured data.

# 4 Modeling setup and hypotheses

The modeling setup of this thesis is presented here and described in more detail in Sections 5 and 6. Figure 1 represents the modeling setup. First, the carbon stocks of the experiment sites are modeled. This is done using a mathematical model that describes the carbon content of soil at different depths, and the measurements of carbon concentration collected as part of the Carbon Action experiment. As the basic sites lack data from deeper soils, the collective data from all the sites is used to define the general vertical distribution of carbon in soils. Bayesian methods are used in modeling: the model acts as the prior, and the posterior is formed by conditioning the model on measured data using Markov Chain Monte Carlo (MCMC). As a result of the modeling, we have estimates of the one-meter carbon stocks of each site before and after the experiment.

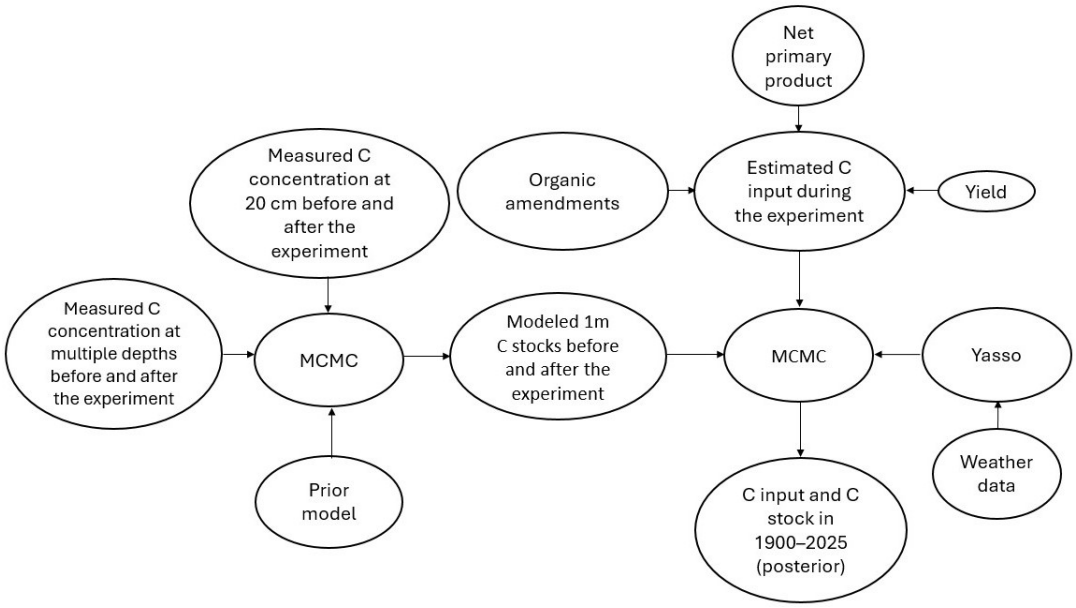



**Figure 1:** A schematic describing the modeling.

Observed carbon inputs during the experiment are estimated as well. For this, we use data of organic amendments added to sites and yields harvested from the sites, and estimated NPP on the sites. The carbon inputs are not calculated for all the years and all the sites, but for those years that we have data of yields.

Next, the long-term carbon stocks and carbon inputs of the sites are simulated. For this, we use Yasso with cropland parameters and weather data from FMI, and prior distributions for initial carbon stocks and carbon inputs. The simulation is done for years 1900-2025. The posteriors are again formed with Bayesian methods: the parameters of Yasso are conditioned with the observations of carbon stocks and carbon inputs, using MCMC for sampling. At each step of MCMC, Yasso is run with the current parameters, and the results are evaluated using observed carbon stocks and carbon inputs.

Once the long-term carbon stocks and carbon inputs have been simulated, they will be examined for the years covered by the experiment (2019–2023). Additionally, we will examine the GPP on the sites. Sites using cover crops or nutrient-rich additions as carbon farming method are considered especially interesting: The mechanisms of these methods are based on increasing photosynthesis or carbon inputs into the soil, which can both be detected with our methods. In addition, they were among the most popular methods in the Carbon Action experiment. Therefore, they are observed as their own groups.

## 4.1 Hypotheses

While the carbon stocks of agricultural soils are declining, using carbon farming practices is expected to at least reduce carbon emissions from the soil. In the Carbon Action experiment, sites using organic amendments had the highest planned carbon inputs, while sites using cover crops were among the plans with lowest carbon inputs (Mattila et al., 2022). Using cover crops, on the other hand, can increase photosynthesis of a field.

Therefore, our hypotheses are the following: 1) The stocks of the treatment sites decrease less than the stocks of the control sites. 2) N-rich treatment sites show the greatest increase in carbon input, in comparison to their control site. 3) Treatment sites using cover crops have a higher GPP than their control sites.

# 5 Modeling one-meter carbon stock

In order to evaluate carbon stock of the soil, more data than only surface level soil samples are needed. As most of the experimental fields have only one sample from near the surface and lack data from the deeper layers, rest of the data is modeled. The modeled data is later used as an input when long-term stocks are simulated. The modeling is done using Bayesian inference: the prior distribution is formed using assumptions of the SOC distribution in the soil, and the posterior distribution is formed by updating the prior distribution with the measured soil carbon concentrations.

The carbon concentration is modeled as a function of the cumulative mass of the soil per area as we move downward. The decreasing nature of SOC with respect to depth in mineral soil can be seen in Figure 2, which represents the measured carbon concentrations of the advanced fields. In this thesis, the prior distribution of SOC is modeled to decrease exponentially as a function of the mass.

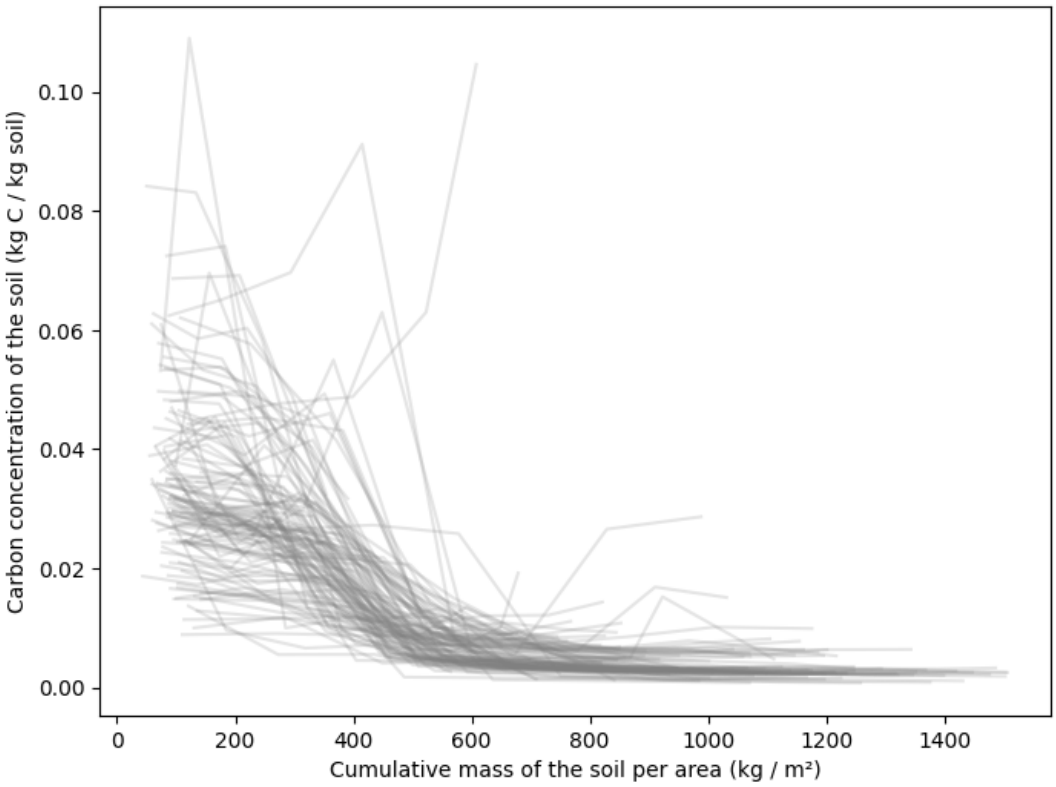



**Figure 2:** The carbon concentration of all of the advanced fields, plotted as a function of the cumulative mass of the soil.

As Figure 2 shows, the carbon contents do not always follow an exponentially decreasing curve. Instead, it seems that, for many fields, the carbon content stays in similar values for the first few data points and then drastically decreases. However, the exponential model was still chosen for a few reasons. First, an exponential model is quite simple, always positive, and fits the prior information of the decreasing nature of SOC concentration. Second, there are many shapes of carbon content profiles,

including exponentially behaving ones. Estimating the shape of a profile with only one data point could be impossible. Third, we are not exactly interested in the distribution of carbon, but of the total amount of carbon to one meter depth.

## 5.1 Prior

The symbols used to form the prior distribution are gathered in Table 1. The prior distribution is built as follows.

**Table 1:** Variables and parameters used in the model, their meaning and the unit. Some are defined with a value, some with distribution, some with a formula.

| Symbol | Description | Formula/distribution | Unit |
|---|---|---|---|
| $c_0$ | Carbon concentration (C%) at 0 cm depth | $N(\mu, \sigma^2)$ | $\frac{\text{kg C}}{\text{kg soil}}$ |
| $\mu$ | Mean of C% at 0 cm | $U(0, 0.5^2)$ | - |
| $\sigma$ | Standard deviation of C% at 0 cm | $U(0, 0.5^2)$ | - |
| $M$ | Cumulative mass of the soil | - | $\text{kg/m}^2$ |
| $M_a$ | An estimated mass of soil at the depth of one meter | 1500 | $\text{kg/m}^2$ |
| $\bar{\gamma}$ | Average rate of exponential decrease | $\log(\bar{\gamma}) \sim N(0, 1)$ | - |
| $\gamma_i$ | Rate of exponential decrease of profile $i$ | $\log(\gamma_i) = \log(\bar{\gamma}) + \log(\delta_i)$ | - |
| $\delta_i$ | Deviation of $\gamma_i$ | $N(0, \tau^2)$ | - |
| $\tau$ | Standard deviation of $\delta$, describes how similar the profiles are | $U(0, 1)$ | - |
| $c(M)$ | C% at depth $M$ | $c_0 e^{-\frac{M}{\gamma M_a}}$ | $\frac{\text{kg C}}{\text{kg soil}}$ |
| $C(M)$ | Cumulative carbon stock at depth $M$ | $\gamma M_a c_0 (1 - e^{-\frac{M}{\gamma M_a}})$ | $\frac{\text{kg C}}{\text{m}^2}$ |

Let the SOC concentration of the soil at soil surface be denoted by $c_0 \in \mathbb{R}^n$, $c_0 = [c_{0,1}, ..., c_{0,n}]$, with $n$ as the number of soil profiles. It is assumed that $c_0$ is normally distributed,

$$c_0 \sim N(\mu, \sigma^2),$$

with normally distributed mean $\mu \in \mathbb{R}$ and standard deviation $\sigma \in \mathbb{R}$,

$$\mu \sim U(0, 0.5^2),$$
$$\sigma \sim U(0, 0.5^2).$$

As stated in 2.2, the SOC concentration of a mineral soil is between 0 and 10%, and often less than 5%, even in the topsoil. The prior distribution $N(0, 0.5^2)$ therefore

gives a very weakly informative distribution of $c_0$. However, this is a deliberate choice: our intention is to steer the final distribution with our initial guesses as little as we can, in order to give more importance for the data.

Next, let the rate of exponential decrease of SOC be defined by $\gamma \in \mathbb{R}^n$, $\gamma = [\gamma_1, ..., \gamma_n]^T$: the smaller $\gamma_i$, the faster the SOC of profile $i$ decreases. It is assumed that the shapes of the profiles are somewhat similar, even though they have their own $\gamma_i$'s. Therefore, for each $\gamma_i$ it holds that

$$\log(\gamma_i) = \log(\bar{\gamma}) + \log(\delta_i),$$

where $\bar{\gamma}$ is the mean of $\gamma$, and $\log(\delta_i)$ the profile-specific deviation of $\log(\gamma_i)$ from $\log(\bar{\gamma})$. The logarithm of $\bar{\gamma}$ follows a normal distribution,

$$\log(\bar{\gamma}) \sim N(0, 1).$$

The deviation is also normally distributed, and the parameter $\tau$ determines how much variation there is between the profiles:

$$\begin{aligned} \log(\delta_i) &\sim N(0, \tau^2), \\ \tau &\sim U(0, 1). \end{aligned}$$

As the prior distribution is formed, the parameters $\mu$, $\sigma$, $c_0$, $\delta$, $\tau$, $\bar{\gamma}$ and $\gamma$ are sampled according to their distributions. Next, these parameters are used to calculate the prior predictive of the SOC stock.

Let $M$ denote the cumulative soil mass per area ($\mathrm{kg/m^2}$) as we move downward:

$$M = \int_0^z \rho(z)dz, \tag{1}$$

where $z \in \mathbb{R}$ denotes the depth in meters and $\rho(z)$ the bulk density of the soil in $\frac{\mathrm{kg}}{\mathrm{m}^3}$. Let $c(M)$ denote the SOC concentration at the corresponding depth, with kg C/kg soil as a unit. Now, let us assume that the SOC concentration decays exponentially as a function of the soil mass $M$ and with the rate of exponential decrease $\gamma$,

$$c(M) = c_0 e^{-\frac{M}{\gamma M_a}}, \tag{2}$$

where $M_a = 1500\mathrm{kg/m^2}$ is the typical soil mass per area, at the depth of one meter, for soil with a low SOC content. By integrating (2) we get the cumulative SOC stock at depth $M$,

$$\begin{aligned} C(M) &= \int_0^M c_0 e^{-\frac{m}{\gamma M_a}} dm \qquad (3) \\ &= \gamma M_a c_0 (1 - e^{-\frac{M}{\gamma M_a}}). \qquad (4) \end{aligned}$$

Equations (2) and (4) can also be expressed profile-wise,

$$c_i(M) = c_{0,i} e^{-\frac{M}{\gamma_i M_a}}, \tag{5}$$

$$C_i(M) = \gamma_i M_a c_{0,i} (1 - e^{-\frac{M}{\gamma_i M_a}}). \tag{6}$$

As there is data from multiple layers, the carbon stocks are also modeled in layers, so that the modeled values can be compared to the measurements. Let $M_{ij,1} \in \mathbb{R}$ denote the cumulative mass at the bottom of the layer $j \in \{1, ..., k_i\}$ in profile $i$, and $M_{ij,0} \in \mathbb{R}$ the cumulative mass at the top of the same layer (therefore, $M_{ij,1} \geq M_{ij,0}$). From the basic fields, there is only data from one layer, so $k_i = 1$. For advanced fields, $k_i \in \{4, ..., 11\}$, as there are data from 4-11 layers. The SOC stock of layer $j$ in profile $i$ can be presented as the difference between observed cumulative SOC stocks at the bottom and the top of the layer:

$$\begin{aligned} C(M_{ij,1}) - C(M_{ij,0}) &= \int_{M_{ij,0}}^{M_{ij,1}} c_{0,i} e^{-\frac{m}{\gamma_i M_a}} \, dm \\ &= \gamma_i M_a c_{0,i} (e^{-\frac{M_{ij,0}}{\gamma_i M_a}} - e^{-\frac{M_{ij,1}}{\gamma_i M_a}}). \end{aligned}$$

Now, let $Y_{ij}$ denote the observed carbon content of the layer $j$ in the $i^{\text{th}}$ profile, and $\bar{Y}_{ij}$ the corresponding real carbon content. Our assumption is that the real carbon content of the layer follows our model,

$$\begin{aligned} \bar{Y}_{ij} &= \frac{C(M_{ij,1}) - C(M_{ij,0})}{M_{ij,1} - M_{ij,0}} \\ &= \frac{\gamma_i M_a c_{0,i} (e^{-\frac{M_{ij,0}}{\gamma_i M_a}} - e^{-\frac{M_{ij,1}}{\gamma_i M_a}})}{M_{ij,1} - M_{ij,0}}, \end{aligned}$$

while the measured carbon content contains observation errors,

$$Y_{ij} = \frac{\gamma_i M_a c_{0,i} (e^{-\frac{M_{ij,0}}{\gamma_i M_a}} - e^{-\frac{M_{ij,1}}{\gamma_i M_a}})}{M_{ij,1} - M_{ij,0}} + \varepsilon_{ij}. \tag{7}$$

Here, $\varepsilon_{ij} \in \mathbb{R}$ represents the observation error. It is assumed to be normally distributed with a half-normally distributed standard deviation,

$$\begin{aligned} \varepsilon_{ij} &\sim N(0, \sigma_\varepsilon^2) \\ \sigma_\varepsilon &\sim N^+(1). \end{aligned}$$

As our observations, we use the data obtained with Equation (7). The data is used to estimate the likelihood of our model, and to update the parameter values to form the posterior distribution.

## 5.2 Likelihood

We are interested to find out what the distribution of observed carbon content $Y_{ij}$ is, and how likely our model would produce results equal to the measured data, given the current parameter values. Let the uncertain parameters of our model be denoted by $\theta = \{\theta_1, ..., \theta_n\}$, where $\theta_i = [c_{0,i}, \gamma_i, \delta_i, \mu, \sigma, \bar{\gamma}, \tau]^T$. Now, we examine the likelihood

$$p(Y_{ij} \mid \theta_i),$$

and want to find out what its distribution is. From Equation (7), we have

$$Y_{ij} = \bar{Y}_{ij} + \varepsilon_{ij}. \tag{8}$$

As $\bar{Y}_{ij}$ is only dependent on $\theta_i$, it is constant with our current parameter values. This means that the distribution of $Y_{ij}$ follows the shape of the distribution of $\varepsilon_{ij}$, but its mean is displaced by $\bar{Y}_{ij}$:

$$Y_{ij} \mid \theta_i \sim N(0, \sigma_\varepsilon^2) + \bar{Y}_{ij} \tag{9}$$

$$\Rightarrow Y_{ij} \mid \theta_i \sim N(\bar{Y}_{ij}, \sigma_\varepsilon^2). \tag{10}$$

Now, (10) is the distribution of our model, given $\theta_i$. Therefore, the likelihood function is the probability density function of the normal distribution $N(\bar{Y}_{ij}, \sigma_\varepsilon)$:

$$p(y \mid \theta_i) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon}} e^{-\frac{(y-\bar{Y}_{ij})^2}{2\sigma_\varepsilon^2}}, \tag{11}$$

where $y \in \mathbb{R}$ denotes a random value of the carbon content.

## 5.3 Posterior

Finally, we can form the posterior distribution of our model. We are looking for an answer to the question: "if we have this measured data $y$, what are the most likely values for our parameters, and how are the parameters distributed?". That is, we want to find out the distribution $p(\theta \mid y)$. Now, $y$ denotes the measured data of the carbon content.

Using Bayes' theorem, we have

$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)}. \tag{12}$$

Here, $y$ is fixed, as it is our data, and $\theta$ is our variable. Therefore, $p(y)$ is a constant and we can express $p(\theta \mid y)$ through proportionality:

$$p(\theta \mid y) \propto p(y \mid \theta)p(\theta).$$

From Equation (11), we get $p(y \mid \theta)$. The distribution of $p(\theta)$, on the other hand, is more complicated. We know the distributions of each variable, but their joint distribution is analytically very heavy to calculate. Therefore, instead of analytically calculating the joint distribution, we will determine the distribution numerically. This is done using a Markov Chain Monte Carlo (MCMC) algorithm No U-Turn Sampler (NUTS) (Hoffman et al., 2014).

The MCMC algorithm is given an initial guess of $\theta$, which is our prior distribution. Using data $y$, the algorithm will try different values for $\theta$, and evaluate whether a new value is better than the current one. The algorithm goes through two chains of 10 000 iterations. The first 5000 iterations are used as a warm-up for the algorithm, and therefore discarded. As the end of the algorithm is reached, we are left with two sets

of 5000 samples of each carbon stock. The sets are evaluated using the Gelman-Rubin statistic (Gelman and Rubin, 1992). The last 500 samples of both chains are chosen to form our final distribution. In one sample, all the sites have a shared value for the parameter $\mu, \sigma, \bar{\gamma}$, and $\tau$, while each site has its own $c_{0,i}, \gamma_i$ and $\delta_i$ in each sample. The 1000 samples now describe the posterior distribution.

## 5.4 Estimating mass at one meter depth

In order to model carbon as a function of cumulative mass of soil at one meter depth, we need to have a value for the cumulative mass. For most fields, we do not have such data. The masses are therefore estimated from the data that we have.

By inverting Equation (1) to represent depth $z^* = 1\text{m}$ as a function of cumulative soil mass $M$, we have

$$z^* = \int_0^{M^*} \frac{dM}{\rho}, \tag{13}$$

where $M^* \in \mathbb{R}$ is the cumulative mass at one meter depth. Bulk density can be estimated using an equation from Heikkinen et al. (2013):

$$\rho = \phi - \omega \log(c), \tag{14}$$

where $\phi = 1.52$, $\omega = 0.280$, and $c$ is the carbon content in percentages. We have an equation for carbon content as a function of cumulative mass in Equation (2). This and (14) can be placed in (13), which gives us:

$$z^* = \int_0^{M^*} \frac{dM}{\phi - \omega \log(c_0 e^{-\frac{M}{\gamma M_a}})},$$

From this, $M^*$ can be solved:

$$M^* = \frac{(\phi - \omega \log(c_0))\gamma M_a}{\omega}(e^{\frac{\omega z^*}{\gamma M_a}} - 1). \tag{15}$$

The evaluation of mass at one meter is done simultaneously with sampling: at each iteration, $M^*$ is calculated again, using the current $\gamma$ in Equation (15).

## 5.5 Validation of the modeled profiles

The results of the soil profile modeling are validated with five-fold cross validation, using data from the advanced farms. The model is run with only the advanced farms, using most of them for training the model and six of them for testing. The validation is run five times, and in each time the test farms are randomly selected. The different divisions of data into test and train are called folds.

In the validation run, the model is given data from all of the layers of the train sites, but for test sites, only the data from the top layers. This imitates the real run where the sites of basic farms only have data from depth of 20 cm. The cumulative carbon stocks

are modeled to the depth of the deepest measurement of each site. After the run, the real values of the test sites can be compared to the values calculated by the model.

The validation is carried out on a farm-by-farm basis rather than on a site-by-site basis in order to reduce data leakage: The sites of one farm are likely to be similar to each other. Using measurements from the same farm as both training data and test data would give the model a benefit that it will not have when the modeling is done for all sites.

# 6 Simulating the long term carbon stock

Next, the long term carbon stocks of the sites are simulated. The simulation is done similarly as the estimation of the one-meter carbon stocks: First, Yasso is given prior distributions of the initial carbon stocks and carbon inputs, and the weather data. Using these, Yasso simulates the carbon stocks for the years 1900-2025. Then, the data that we have of the carbon input (NPP, organic amendments, yields), and the modeled one-meter carbon stocks, are used to form posterior distributions of simulated carbon inputs and initial carbon stocks. These distributions are then further used to simulate the posteriors of the long-term carbon stocks.

In order to simulate the carbon stocks, some assumptions are made about the system. We assume that the areas were originally covered by forests that were old enough to have reached a steady state in terms of carbon: the SOC stock of the soil was constant, as was the carbon input (litterfall) into the soil. Decomposition of carbon was constant as well, determined by the average climate condition. At the start of the simulation period in year 1900, the land was converted into a cropland. After conversion, we assume that the carbon input has varied, depending on the weather and cultivation practices. Similarly, we assume that the rate of carbon decomposition varies between years. However, we assume that the composition of the carbon input and the decomposition of carbon into different pools remain constant.

## 6.1 The model

Let us start by defining the state variable $\mathbf{x}(t) = [x_1(t), ..., x_{15}(t)]^{\mathrm{T}}$. The state variable $\mathbf{x}(t)$ represents the mass of carbon stock in the soil at timestep $t$ to the depth of 1 meter. The carbon stock is divided into 15 pools, each represented by one element of $\mathbf{x}(t)$: five pools based on the solubility of the carbon, and each pool further divided into three size classes. Let us denote the pools with $k \in \{1, ..., 15\}$. The net carbon input into the system (litter from plants and organic amendments) is denoted by $\mathbf{b}(t) = [b_1(t), ..., b_{15}(t)]^{\mathrm{T}}$. Likewise, each element b(t) tells us the mass of carbon entering the pool $k$ at time $t$.

Let $\mathbf{A}(t) \in \mathbb{R}^{15\times15}$ be the transfer of carbon between the pools. Each element represents the fraction of carbon decomposing from one pool into another. The difference between decomposing carbon and carbon entering other pools is the carbon being released as $CO_2$. The distribution of carbon from one pool into the others is constant for the given pool, but how much decomposes per time step depends on the conditions (temperature and precipitation). Similarly, the proportion of decomposing carbon that is converted into $CO_2$ is constant for each pool. The transfer of carbon $\mathbf{A}(t)$ is constructed similarly as in Viskari et al. (2022), where a more precise description of the matrix and its elements can be found.

The change in the carbon stocks at time $t$, or, our state equation, is

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{b}(t). \tag{16}$$

Now, based on our assumptions, the land was taken into agricultural use at time $t = 0$, and before the conversion, there had been a steady-state forest on the land:

$$\mathbf{x}_0 = \lim_{t\to\infty}\mathbf{x}(t),$$

where $\mathbf{x}_0 = \mathbf{x}(0)$. We can express the carbon input of a steady state forest as

$$\mathbf{b}_f = \alpha_0\mathbf{u}_f, \tag{17}$$

where $\alpha_0 \in \mathbb{R}$ is the total steady state carbon input, and $\mathbf{u}_f \in \mathbb{R}^{15}$ partitions the input into fractions that are typical for a forest litter. Following Viskari et al. (2022), in a steady state, the carbon stock becomes

$$\lim_{t\to\infty}\mathbf{x}(t) = -\mathbf{A}^{-1}\mathbf{b},$$

where $\mathbf{A}$ and $\mathbf{b}$ are the constant, steady state carbon transfer and input. Due to our assumptions of $\mathbf{x}_0$, we can express it as

$$\mathbf{x}_0 = -\alpha_0\mathbf{A}_f^{-1}\mathbf{u}_f,$$

where $\mathbf{A}_f$ is the average transfer of carbon between the pools in the steady-state forest, determined by the average climate conditions between years 1961-1990.

After the site has been converted into a cropland, the carbon input is no longer constant, but the relative distribution of carbon input into the different pools is constant. Here, the distribution is determined by $\mathbf{u}_c \in \mathbb{R}^{15}$, which describes the typical composition of cropland biomass inputs. The total carbon input is divided into three components:

$$\mathbf{b}(t) = (\beta_0 + \beta(t) + \delta(t))\mathbf{u}_c, \tag{18}$$

where $\beta_0$ represents the constant carbon input, $\beta(t)$ is the carbon input that varies on a long, decadal scale, and $\delta(t)$ the yearly variation in the carbon input.

We assume that the weather conditions and carbon inputs stay constant in each time step. Now, $\mathbf{x}(t)$ can be calculated one time step at a time, starting from $\mathbf{x}_0$ and using Equation (16).

## 6.2 Prior distributions

In the model, the distribution of carbon between different pools is predetermined: partitions of carbon input $\mathbf{u}_f$ and $\mathbf{u}_c$ are constant, and only the rate of decomposition changes, not how carbon transfers between the pools. These will stay the same regardless of the measured data.

What changes, in turn, is the value of $x_0$, and the magnitude of carbon inputs before and after conversion into cropland. These are given prior distributions, which are later updated using the data.

The initial carbon stock $\mathbf{x}_0$ is dependent on the total steady state carbon input $\alpha_0$, which is considered to be normally distributed:

$$\alpha_0 \sim N(\mu_\alpha, \sigma_\alpha^2).$$

The prior value for $\mu_\alpha$ is chosen such that the expected value of the total initial carbon stock is 15 kg C / m$^2$, according to an estimation by Akujärvi et al. (2014):

$$E[\sum_{i=1}^{15} x_{0_k}] = 15 \text{ kg C / m}^2.$$

The standard deviation $\sigma_\alpha$ is given a prior value of 5 kg C / m$^2$.

In the prior, the constant carbon input $\beta_0$ follows a gamma distribution. Its scale and rate are chosen such that its value is between 0.2 and 1.5 kg C / m$^2$/ year with a probability of 95%. The slowly changing carbon input $\beta(t)$ is a Gaussian process that satisfies the following conditions:

$$\beta(t) \sim N(0, r^2)$$
$$Cov(\beta(t), \beta(t')) = r^2 e^{-\frac{(t-t')^2}{2s^2}},$$

where $r = 0.2$ kg C / m$^2$/ year and $s = 30$ years. The yearly variation is considered to be normally distributed:

$$\delta(t) \sim N(0, r^2).$$

## 6.3 Conditioning the model with data

Now, we have formed the prior. Next, we use data of the carbon stocks, NPP, yields and organic amendments to form the posterior of our model. The modeled carbon stocks of one meter, calculated in section 5 with Equation (6) using measurements of the carbon concentration, are used as observations of $\mathbf{x}$:

$$\sum_{k=1}^{15} x_{ik}(t_c) = C_i(M_m) + \varepsilon_{c_i}. \tag{19}$$

Here, $x_{ik}$ is the carbon stock of pool $k$ in site $i$, $t_c$ is the time at which the carbon concentration measurement or measurements were taken, and $M_m$ is the mass of soils at the depth of one meter. The observation error $\varepsilon_{c_i}$ is assumed to follow a zero-mean normal distribution. The variance of the error is the same as the variance of the posterior distribution of profile $i$.

Observations of the total carbon input $b$ are calculated using the farming data of the yields and organic amendments, and satellite data of the NPP: The carbon entering the system is the biomass from plant growth (NPP) $B_{npp}$ and biomass brought to the field (organic amendments) $B_a$, minus the biomass taken away from the field (the yield) $B_y$:

$$\sum_{k=1}^{15} b_{ik}(t_b) = B_{npp} + B_a - B_y + \varepsilon_{b_i}. \tag{20}$$

The observation error $\varepsilon_{b_i}$ is again normally distributed with mean of zero. The variance is the same as the variance of NPP: for yield and organic amendments, we have no estimates of the error. We can now see that (19) and (20) are of the same form as (8). Therefore, we can calculate their likelihoods similarly as in (11). Furthermore, the posteriors can be calculated using the Bayes' theorem, as we did in 5.3. Again, MCMC algorithm NUTS is be used for sampling the posterior.

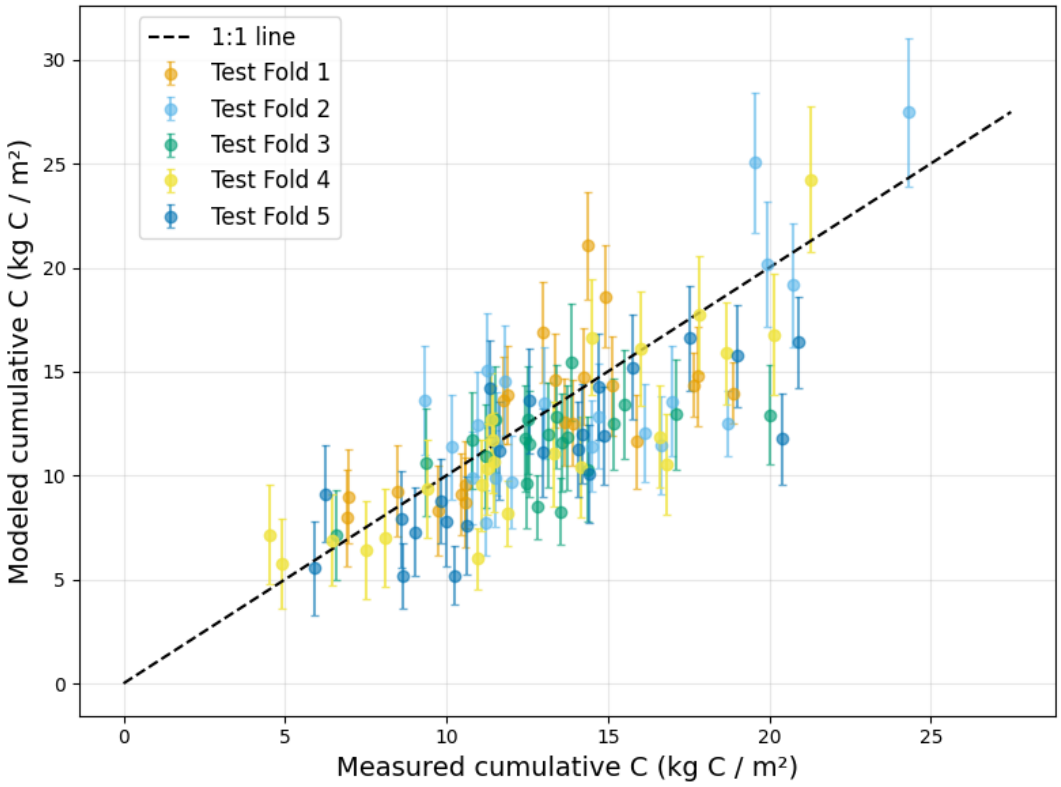



**Figure 3:** Modeled cumulative carbon of the test sites plotted against the measured values of cumulative C. A dot indicates the mean of the MCMC generated samples, while an error bar indicates the standard deviation.

# 7 Results

## 7.1 Modeled one-meter carbon stocks

Let us start by observing results of the five-fold cross-validation of the modeling of one-meter carbon stocks (described in Section 5.5). Figure 3 shows the cumulative carbon stocks of the test sites of all five folds. The cumulative carbon of each test site has been modeled to the deepest measured layer on the given site. The measured carbon stock of a layer is calculated as the product of the measured mass and carbon concentration of the given layer. The measured cumulative carbon stock is the sum of the carbon stocks of the layers. The results of cross validation look adequate: when all results from all the folds are combined, the coefficient of determination $R^2$ of the whole dataset is 0.42, and the root mean square error is 2.9 kg C / m$^2$.

Now, let us take a look at the modeled one-meter carbon stocks (model described in Section 5). The SOC stocks of some examples of advanced farms can be seen in Figure 4, along with the measurements of these fields. Based on a visual assessment, the model succeeded to mimic the shape of the cumulative carbon stocks quite well, however, this varied between the fields. For example, the measurements of sites CA-HV039-T and CA-HV023-T corresponded well with the model, whereas with

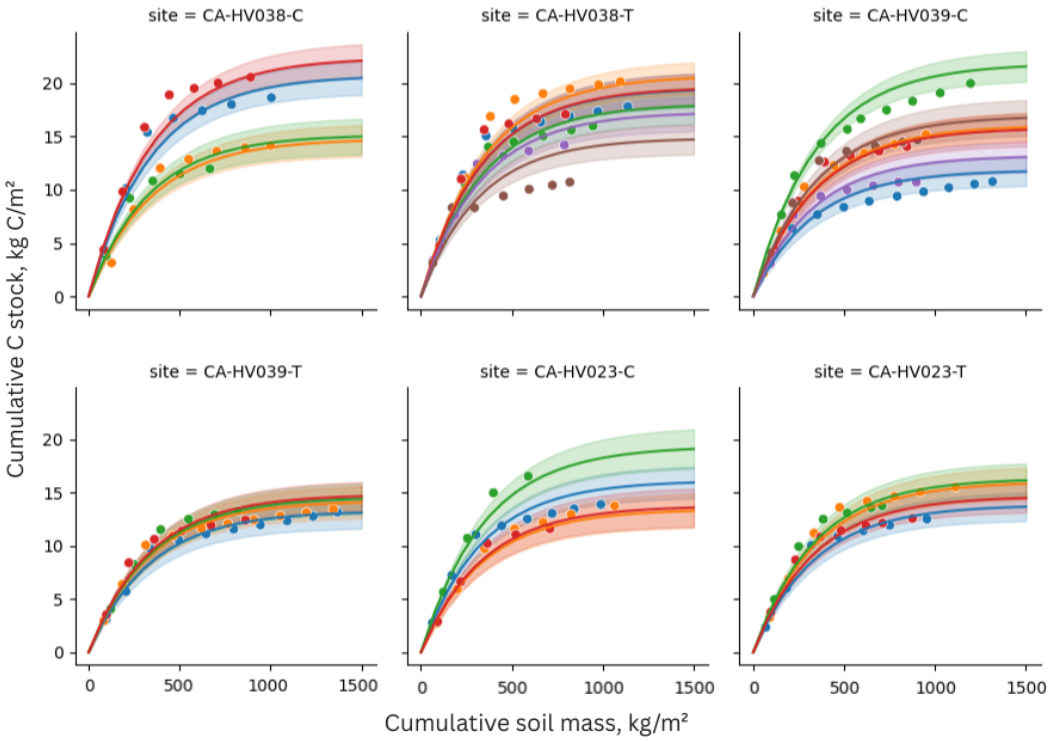



**Figure 4:** The modeled stocks of three advanced farms. The dots represent measurements, and the line is the model. The darker line shows the mean of the modeled samples, and the shading around the model indicates their standard deviation. Each plot is from one experiment site: C at the end of the site name stands for control, while T means treatment. Several sets of measurement have been taken from the same locations, and each plot contains measurements from two years: before and after the experiment period. Each replicate and each year has its own color and its own sampled model parameters.

site CA-HV038-T the modeling was not as successful. It seems that the less there is variation between the measurement replicates, the better the model performs.

Still, we can see that most of the measurements fell within the models' confidence intervals. Furthermore, as the purpose is to model the one-meter carbon stock, it is more important that the estimated value of carbon storage in deeper soil fits the model, than that of the shallow soil. Out of these six example sites, almost all of the deepest measurements fell within the models' confidence intervals.

The results of the modeling of the carbon stocks of a few basic sites can be seen in Figure 5. Similar measurements resulted in similar stocks and overall the stocks resemble each other quite a lot. However, since we have no measurements from deeper in the soils, we cannot say how well the model depicted the actual stocks. The assumption is that modeling of basic farms gives is as accurate as modeling the test sites in five-fold cross validation.

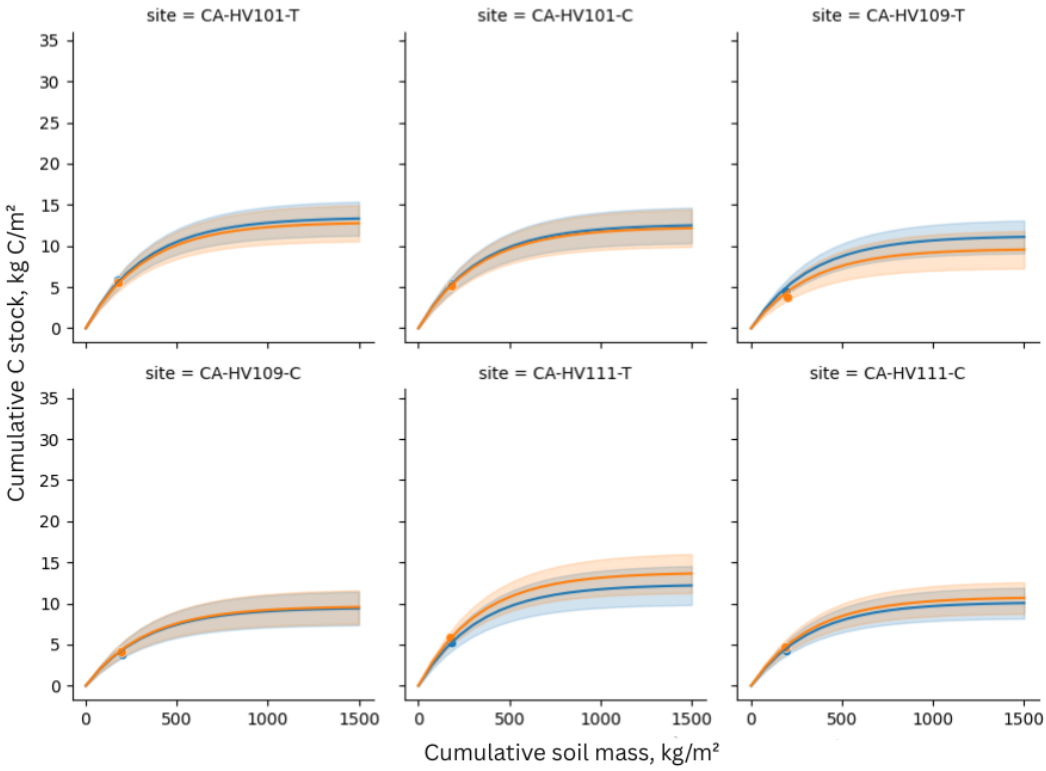



**Figure 5:** The modeled stocks of the basic sites. Each plot is from one experiment site, and contains two measurements: one from before the experiment, and one after.

## 7.2 Development of the stocks over a long term

Next, let us take a look at the development of the carbon stocks over time. Figure 6 shows the simulated carbon stocks and carbon inputs of two farms. Data of carbon stocks before and after the experiment and carbon input during the experiment years are shown in black color. It is good to note that the values of the y-axis differ between the graphs. As we can see, the shapes of all of the carbon stocks look similar, all of them are declining. However, they differ in values. This is due to the logic of the model: the simulation can only adjust the stock to correspond to the measurement points by adjusting the carbon input and the initial value of carbon stock. Only the carbon input can affect the shape of the simulated stock, but the amount of carbon in the inputs, in comparison to the carbon stock, is so small that it cannot change the stock very quickly. We can see that the shape of the simulated carbon input, in turn, is heavily affected by the data: the simulation follows the data of carbon inputs quite closely.

Still, we can see that, while the simulated stock does not exactly correspond to the data points, it still produces similar values. This can be seen particularly when graphs of two different sites are compared to each other: for example, the simulation and data of the stock of site CA-HV023-C are both between 14 and 16 kg C / $m^2$ during years 2018 and 2025, while with CA-HV039-T, their values are between 13 and 15 kg C / $m^2$.

## 7.3 Effects of carbon farming

Let us now take a closer look at the estimated GPP and simulated carbon inputs and stocks during the experiment period, 2019-2023. We start by looking at the results of carbon input and GPP, as these affect the results of carbon stocks. In the carbon input section, the results of nutrient-rich sites are presented separately, as are the results of cover crop sites in the GPP section. Then, we move on to the carbon stocks, where both N-rich sites and cover crop sites are presented as separate groups. Statistics of each of these variables are collected in Tables 2 and 3. In terms of carbon inputs and changes in carbon stocks, the statistics are based on the averages of the samples produced by MCMC: Each sample has an estimate of the carbon stock and carbon input for each site. The sites are grouped based on their carbon farming practice, and whether they are a control or a treatment site. Then, the averages for each sample are calculated from these groups. From these averages, we take the mean and standard deviation of the difference between treatment and control, as well as the 5% and 95% quantiles, which are shown in the tables. Similarly, the statistics of GPP are based on averages of GPP samples given by the photosynthesis estimation method: the sites are first grouped, and then averages of the samples are taken within these groups.

### 7.3.1 Carbon input during the experiment

Figure 7 shows the distributions of the average yearly carbon inputs of the advanced sites between the years 2019 and 2023. The distributions show the 1000 MCMC

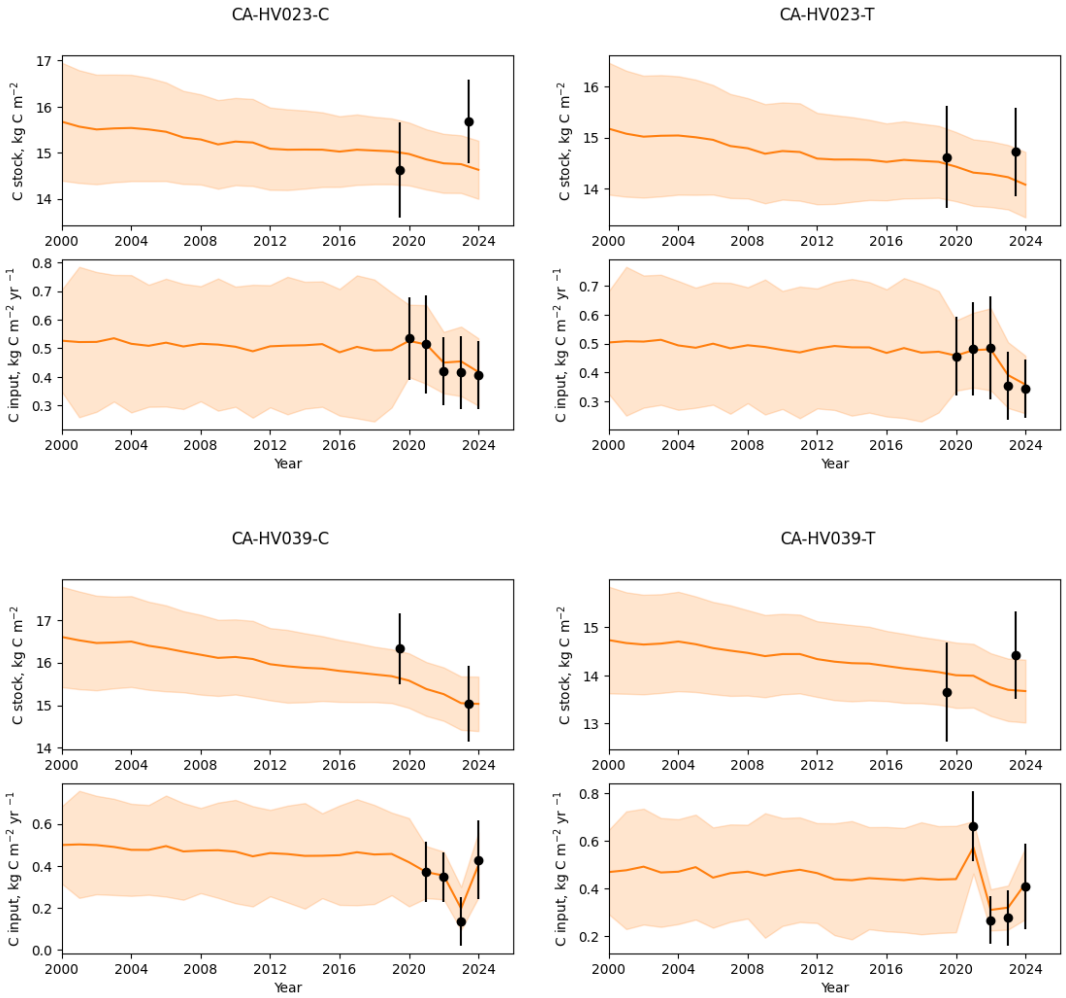



**Figure 6:** Results of the simulation of carbon stocks and inputs from control and treatment sites of two example farms during the last 25 years. In the figures, the upper graph shows the development of the carbon stocks over the years, while the lower graph presents the carbon inputs into the soil. The dark orange line represents the mean of the posterior samples, while the lighter shades show the confidence interval, obtained from the standard deviation of the samples. The black dots in the upper graphs show the estimated carbon stocks before and after the experiment, and the carbon input obtained from estimates of NPP, yields and organic additions in the lower graphs. The standard deviations of both are shown with the black bars.

generated samples. Each sample is the average of the carbon input in the sites in question: sites using nutrient-rich additions as their carbon farming practice, sites using other methods, or all the sites combined. Each of these groups is further divided into control or treatment.

The sites with N-rich additions are separated as their own group here, as we were expecting to observe an increase in the carbon input of these treatment sites in

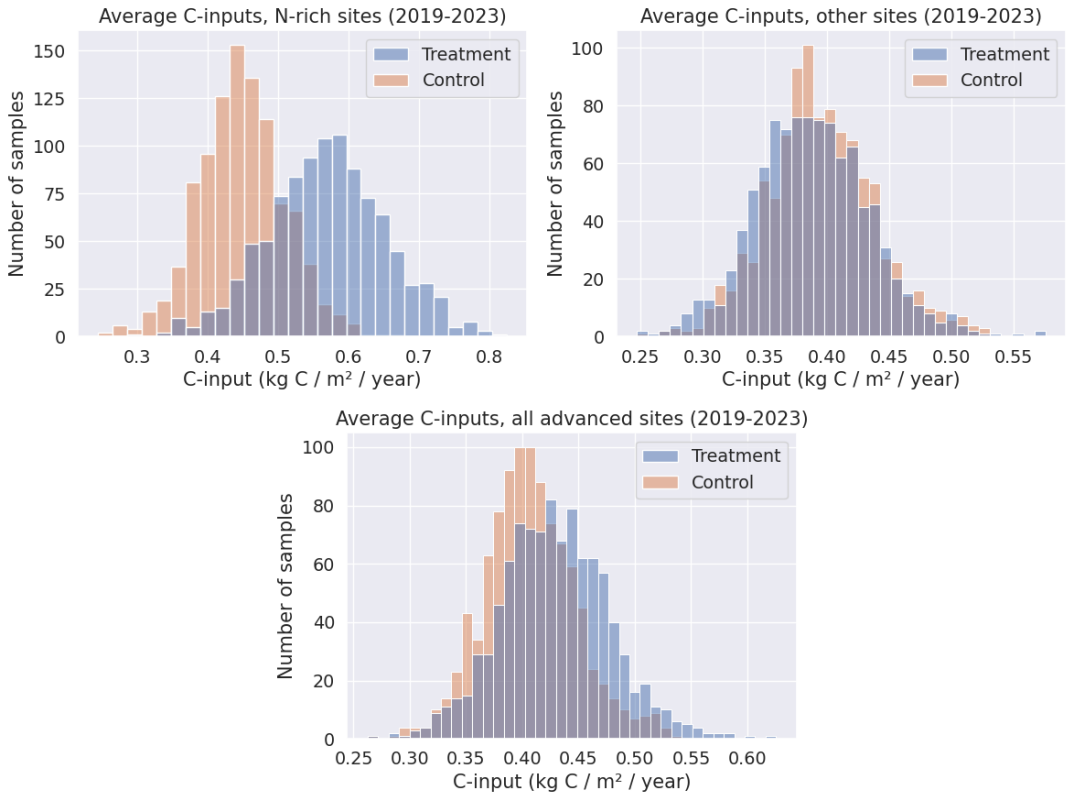



**Figure 7:** Histograms showing the distribution of average carbon inputs of each MCMC sample of the advanced sites. The sites are separated based on the treatment option, and divided into control and treatment. Averaging is done over the sites that belong in the given group, and over the experiment years 2019-2023. Distribution of average carbon inputs of nutrient-rich sites can be seen in the top left corner, and of all the other sites on the top right corner. In the bottom, we have all advanced sites together.

particular. There is indeed a notable difference between the treatment and control sites of N-rich farms. The carbon input of treatment sites was bigger, however, it also had more variation. The carbon input of N-rich sites was also greater than the carbon input of other sites: the average input of N-rich treatment sites is 0.57 kg C / $m^2$ / year, while with other sites (both control and treatment), it is 0.39 kg C / $m^2$ / year. In contrast, with the other treatment options, the control sites actually have had a slightly larger carbon input. However, there is no clear difference to be seen: the average difference is -0.01 kg C / $m^2$ / year, which has a smaller absolute value than the standard deviation of the difference (Table 2). With all the advanced sites together, we can still see that the treatment sites have received more carbon. However, the difference is small and the standard deviation is again larger than the average.

As we extend the scope from advanced sites to all sites and observe the Figure 8, we see similar results. N-rich treatment sites also had a higher carbon input than their control sites. On average, the the treatment sites of N-rich farms had 25% greater carbon input than their control sites. As both 5% and 95% quantiles are positive (Table

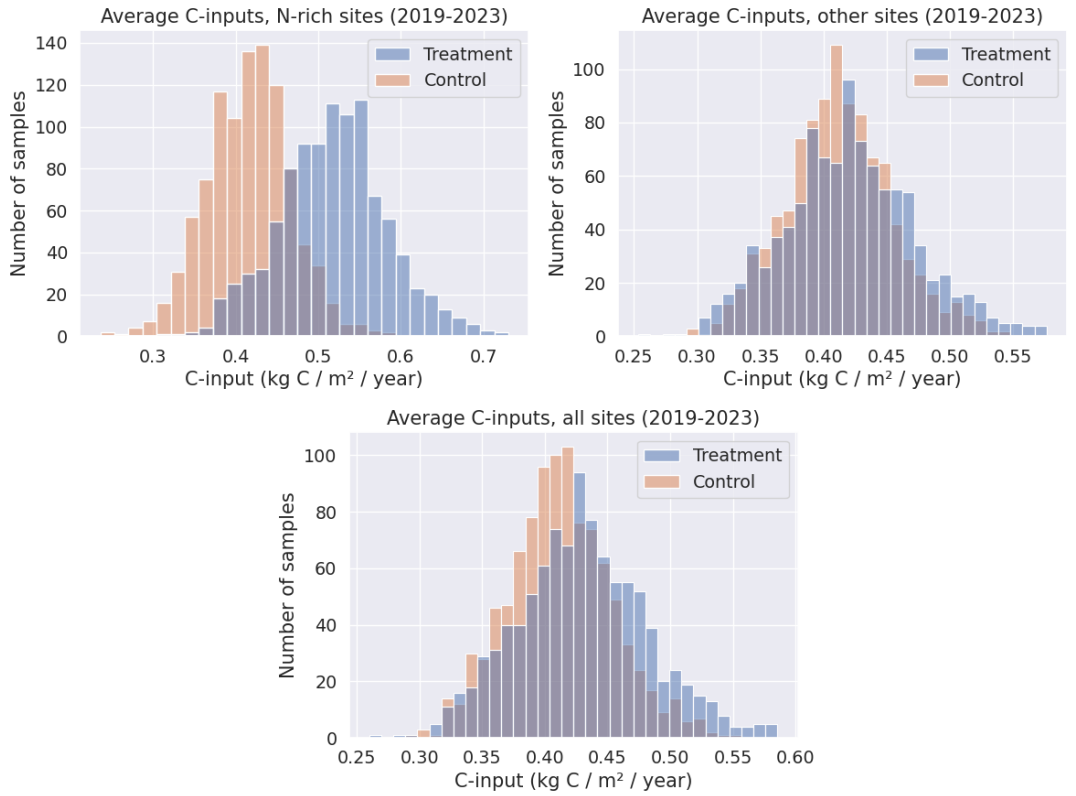



**Figure 8:** Histograms showing the distribution of average carbon inputs of each MCMC sample of all the sites.

3), the difference is significant. Here, unlike with the advanced sites, the treatment sites using other methods had a slightly larger carbon input compared to the control sites. However, the difference is again insignificant, with an average difference smaller than the standard deviation (Table 3). With all sites combined, the results are similar to advanced sites: the treatment sites had a little more carbon inputs, but the difference is small and statistically insignificant.

### 7.3.2 Gross primary production

Next, let us take a look at the gross primary production at the sites during the experiment. Figure 9 shows the means and standard deviations of the sampled annual GPP's of both advanced sites and all the sites. The samples are produced using the photosynthesis estimation method (see Section 3.3). Here, cover crops are raised as their own group, as using cover crops is expected to increase photosynthesis.

In both advanced farms as well as all farms, the sites with cover crops as the carbon farming method have a lower GPP than the sites using other methods. However, the treatment sites with cover crops still have a higher GPP than the corresponding control sites. We can therefore see an improvement in the amount of carbon being sequestered when cover crops are used, as expected. Still, the differences are small and have a large standard deviations in relation to the means (Tables 2-3).

**Table 2:** Statistics of the differences between treatment and control sites (treatment - control) in carbon inputs, GPP, and carbon stock changes. Here, only results of advanced sites are shown.

| | | Treatment - Control | | | |
|---|---|---|---|---|---|
| Variable | Carbon farming practice | mean | std | 5% | 95% |
| C input (kg C / $m^2$ / year) | N-rich | 0.12 | 0.09 | −0.02 | 0.27 |
| | Others | −0.01 | 0.03 | −0.06 | 0.04 |
| | All | 0.02 | 0.03 | −0.03 | 0.08 |
| GPP (g / $m^2$ / year) | Cover crops | 19 | 91 | −110 | 140 |
| | Others | −43 | 74 | −150 | 65 |
| | All | −26 | 61 | −110 | 67 |
| ΔC stocks (kg C / $m^2$ / year) | Cover crops | 0.02 | 0.04 | −0.05 | 0.08 |
| | N-rich | 0.07 | 0.06 | −0.03 | 0.16 |
| | Others | 0.00 | 0.03 | −0.05 | 0.04 |
| | All | 0.02 | 0.03 | −0.03 | 0.06 |

**Table 3:** Statistics of the differences between treatment and control sites (treatment − control) in carbon inputs, GPP, and carbon stock changes. Statistics of all the sites are presented here.

| | | Treatment − Control | | | |
|---|---|---|---|---|---|
| Variable | Carbon farming practice | mean | std | 5% | 95% |
| C input (kg C / $m^2$ / year) | N-rich | 0.1 | 0.04 | 0.03 | 0.18 |
| | Others | 0.01 | 0.02 | −0.02 | 0.04 |
| | All | 0.02 | 0.02 | −0.01 | 0.05 |
| GPP (g / $m^2$ / year) | Cover crops | 15 | 39 | −52 | 71 |
| | Others | 2.1 | 34 | −56 | 58 |
| | All | 6.1 | 27 | −43 | 47 |
| ΔC stocks (kg C / $m^2$ / year) | Cover crops | 0.01 | 0.02 | −0.03 | 0.04 |
| | N-rich | 0.05 | 0.03 | 0.01 | 0.09 |
| | Others | 0.0 | 0.02 | −0.03 | 0.03 |
| | All | 0.01 | 0.02 | −0.02 | 0.04 |

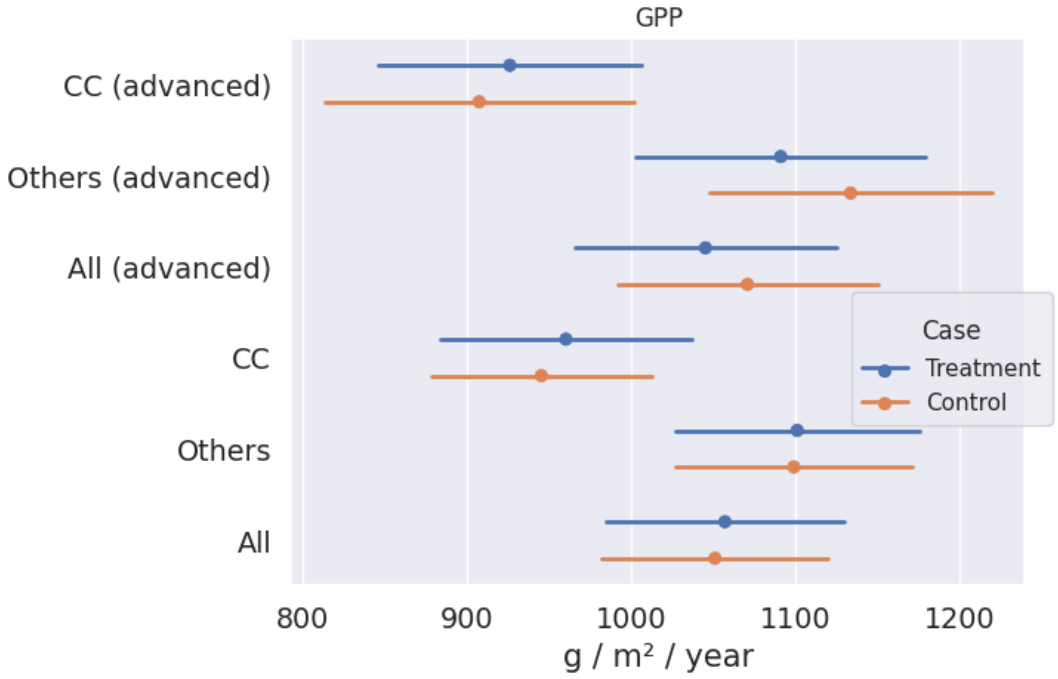



**Figure 9:** A graph showing the average annual GPP on the sites, taken as the mean of the sample averages. The dots represent the mean of the samples, while the lines on each side show the standard deviation.

As we take look at the rest of the advanced sites, we can see that the control sites have actually a higher GPP than the treatment sites. However, if we broaden our scope from advanced sites to all sites, we do not see similar results. This time, the treatment sites are performing better in terms of GPP, however, the difference is very small. Overall, the statistics (Tables 2-3) show that the difference is insignificant: with all the methods, and both advanced sites and all sites, the standard deviations are greater than the absolute values of the average differences. The 5% and 95% quantiles indicate that the differences between control and treatment are fairly evenly distributed between positive and negative values.

### 7.3.3 Development of the carbon stocks during the experiment

Now, let us observe how the carbon stocks have changed during the experiment. Figure 10 shows histograms illustrating the distribution of carbon stock samples of advanced sites in 2019 and 2023. Additionally, we have the distribution of the annual change of stocks in these years, which is simply obtained from the average annual difference between the stocks in 2023 and 2019. At the start of the experiment, the SOC-stocks of treatment sites were on average smaller than those of the control sites: the average stock of treatment sites was 13.01 kg C / $m^2$, while the average stock of control sites was 13.25 kg C / $m^2$. As we observe the stocks in 2023, we notice that the difference between treatment and control sites has decreased, as the average stocks are now 12.78 kg C / $m^2$ and 12.96 kg C / $m^2$, respectively. This can also be seen from the graph showing the change of stocks between these years: Both treatment and control sites have lost carbon, but for the treatment sites, the change has been smaller. The carbon stocks are also illustrated in Figure 11, where we have the mean stocks of all the sites in the years 2019 and 2023. With all sites, we do not see as clear of a difference between the stocks of treatment and control. The treatment sites do have less carbon in 2019, however, the difference is small and possibly caused only by advanced sites. Additionally, there is much more variation with all the sites compared to the advanced sites, which further reduces the significance of the difference.

Changes in stocks of advanced sites can be further observed in Figure 12, where the fields have been divided into three groups: cover crops, nutrient-rich additions, and others. With nutrient-rich additions, we see a clear difference between treatment and control sites: on average, the treatment sites have lost less SOC than the control sites. Similar results can be seen with the cover crop sites, however, the difference is smaller with these sites. With other sites, there is no clear difference. However, the treatment sites seem to have more variation with their change of stock, while the results of control sites are more consistent. As we compare the three graphs with each other, we notice that the sites that have lost the most carbon are the control sites of N-rich and cover crop farms. The average change of stocks in N-rich control sites was $-0.109$ kg C / $m^2$ / year, and in cover crop control sites $-0.097$ kg C / $m^2$ / year. The treatment sites of N-rich farms and the control sites of farms using other practices lost the least carbon: the changes were on average $-0.042$ kg C / $m^2$ / year and $-0.047$ kg C / $m^2$ / year.

The stock changes of all sites (Fig. 13) are similar to the results of advanced sites

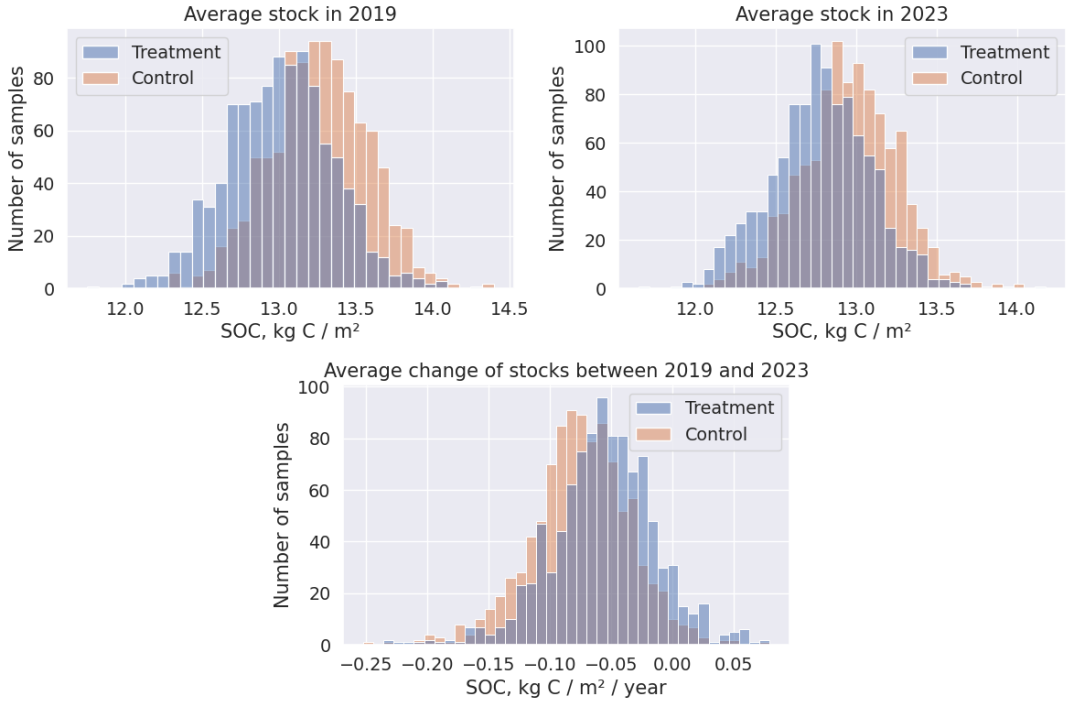



**Figure 10:** Histograms showing the average carbon stock of advanced sites in different samples. The first two graphs show the distributions of carbon stocks in 2019 and 2023. The last figure shows the average annual change in stocks, calculated as the average difference between carbon stocks in these years.

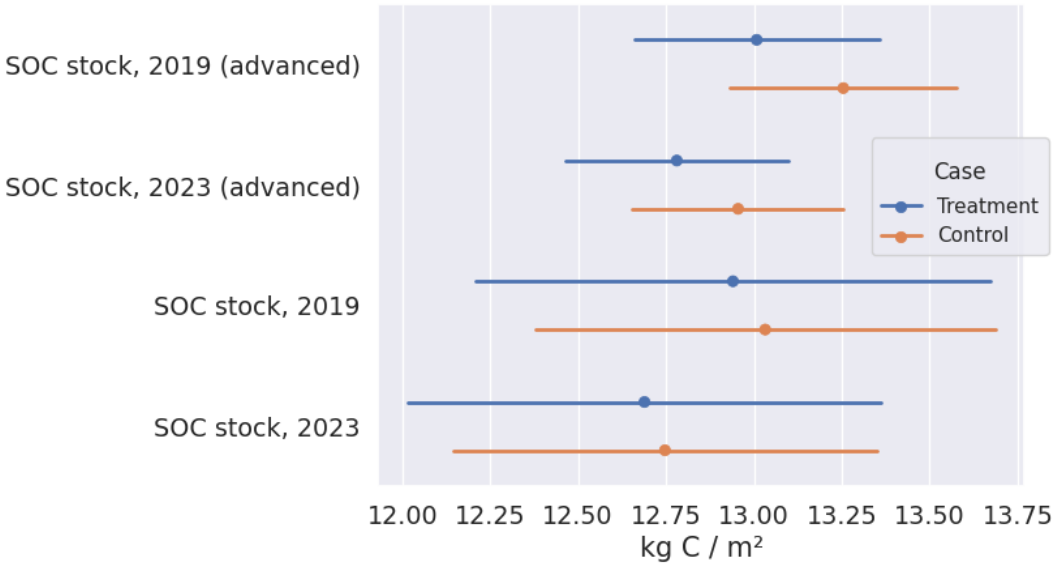



**Figure 11:** Mean and standard deviation of the stocks in years 2019 and 2023.

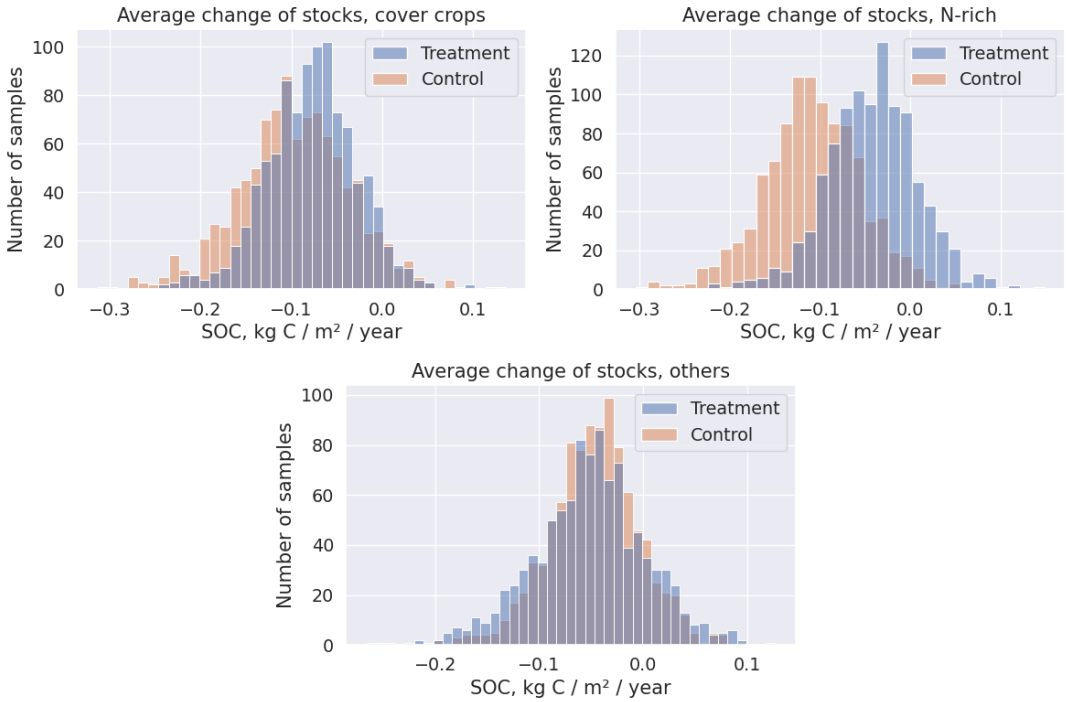



**Figure 12:** Average changes of stocks of advanced sites using cover crops, nutrient-rich additions, and other carbon farming methods. The change is obtained by subtracting the 2019 stock from the stock of 2023.

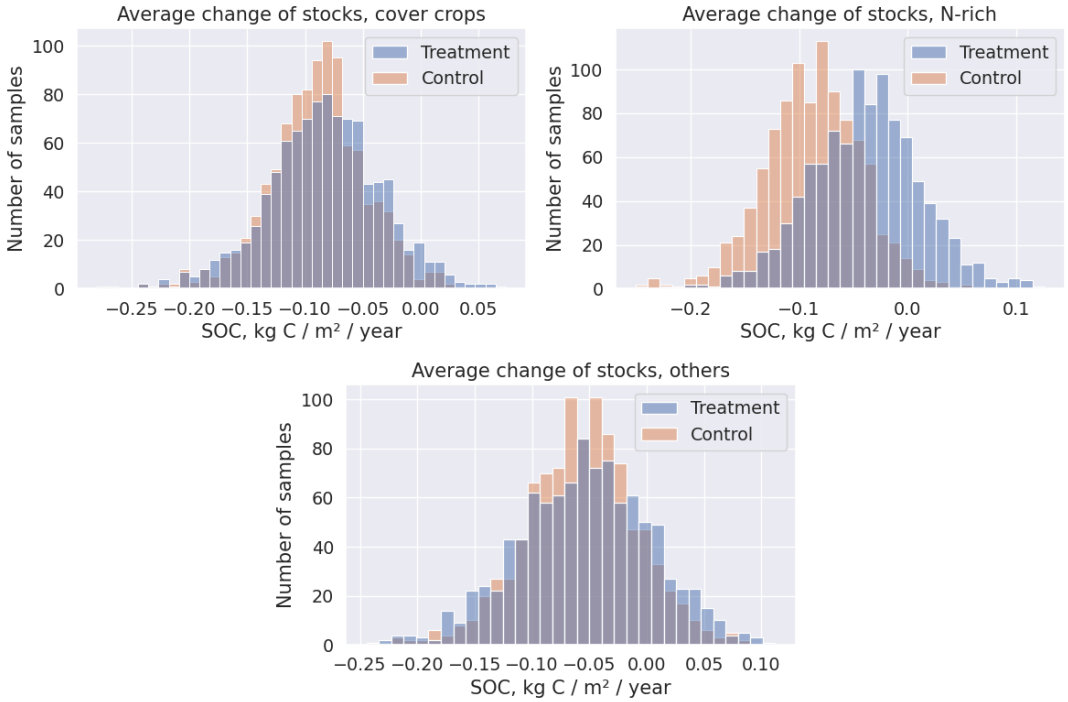



**Figure 13:** Average changes of stocks from all the fields. Divided to sites using cover crops, nutrient-rich additions, and other carbon farming methods.

(Fig. 12): Again, the sites are, on average, losing carbon. N-rich additions have the biggest effect on the change of stock, and sowing cover crops seems to have a small effect as well. With other methods, there is no clear difference, but the treatment sites have more scattered results again. Again, the sites that lost the most carbon are the control sites of N-rich and cover crop farms (−0.088 kg C / $m^2$ / year and −0.090 kg C / $m^2$ / year). The treatment sites of N-rich farms and the treatment sites of farms using other practices lost the least carbon (−0.040 kg C / $m^2$ / year and −0.054 kg C / $m^2$ / year).

Table 4 shows the probabilities that the average carbon loss was smaller in treatment sites than in control sites. Since the probabilities are around 50% in almost all cases, this indicates that there is no significant difference between the treatment and the control. In addition, the results contain a lot of variation, and the differences between treatment and control settle on both sides of zero (Tables 2 and 3). An exception to this are the N-rich sites: both 5% and 95% quantiles of the difference are positive. On average, the treatment sites of N-rich farms lost 55% less carbon than their control sites.

## 7.4 Summary of the results

In summary, modeling of the one-meter stocks of the advanced fields fits the measurements well. As no similar measurements of the basic fields is available, the modeling of basic fields cannot be verified.

**Table 4:** A table showing the probability of change of stock of a random treatment sample being greater than the change of stock of a random control sample.

| Carbon farming practice | $P(\Delta C_T > \Delta C_C)$, advanced sites | $P(\Delta C_T > \Delta C_C)$, all sites |
|---|---|---|
| Cover crops | 0.60 | 0.54 |
| Nutrient-rich additions | 0.82 | 0.76 |
| Other methods | 0.48 | 0.52 |
| All methods | 0.63 | 0.55 |

The long-term simulation of carbon stocks shows a declining trend in the stocks. The simulations do not match the measured data of the stocks perfectly, but they have similar orders of magnitudes. The simulated carbon inputs, on the other hand, adjust to the measured carbon inputs more precisely.

During the experiment, quite a clear difference can be seen in the carbon input of the sites that have been treated with nutrient-rich amendments and their control sites. Our initial hypothesis, that the N-rich sites have received the most carbon in comparison to their control sites (described in Section 4.1), is supported by these results. With other sites, no significant difference between treatment and control can be seen.

In GPP, there is no clear difference between treatment and control sites, with any of the groups. Our hypothesis, that the cover crop treatment sites would have more GPP than their control sites, is therefore not supported. Instead, cover crop farms have had less GPP than the other farms, with both treatment and control sites.

At the start of the experiment, the treatment sites had, on average, less carbon than the control sites. In general the sites have lost carbon, but the treatment sites have lost less carbon than the control sites. This finding supports our first hypothesis (Section 4.1); however, the difference is small. The largest difference in stock change between control and treatment can be found in N-rich farms.

# 8 Discussion

## 8.1 Deeper analysis of the results

Some of the results that we got were expected, as they supported the hypotheses described in Section 4.1. Our second hypotheses, that the N-rich sites would have the biggest difference in carbon input, was supported, and it is easy to see why we see such a big difference with the N-rich farms: More biomass is brought into the treatment sites than the control sites, and this data is given to the simulation. Artificially adding carbon into a field is also quite a secure way of increasing the carbon stock of the soil, as it is not dependent on the crop success, for example.

While our third hypothesis (treatment sites using cover crops have a higher GPP than their control sites) was not supported by the GPP of cover crop sites, the results were not very surprising either: there was a small increase of GPP with the cover crop treatment sites compared to the control sites. Still, the difference is statistically insignificant. The lower GPP of cover crop farms compared to other farms can be explained by the fact that cover crop farms cultivated annual crops. Other farms also had perennial crops as their main crop, which typically have higher photosynthesis rates (Jaikumar et al., 2013).

The treatment sites having on average lower carbon stocks than the control sites at the start of the experiment was not an expected result. As this was especially seen with the advanced sites, it could be only a coincidence: there were only 18 advanced sites, so not many sites are needed to stir the average. However, this raises a question: Did the treatment sites lose less carbon because the carbon farming practices were effective, or because there was less carbon to lose to begin with? Still, it seems that at least the nutrient-rich additions was an effective carbon farming method: The difference between change of stocks of treatment and control is clear, and the treatment sites had a higher carbon input.

The average annual carbon loss from treatment sites was 0.48% of the 2019 carbon stocks, with a standard deviation of 0.41%. From the control sites, the loss was 0.54% with standard deviation of 0.34%. As Heikkinen et al. (2013) estimated the carbon loss from Finnish mineral soils to be 0.4%, we can see that our numbers are a bit elevated. However, the magnitude of the change is similar.

Considering the target of increasing the world's soil carbon stocks by 0.4% annually, this is an undesired result. Even at their best, our carbon farming practices only succeeded to decrease the loss of carbon, and there is still a long way to actually sequester carbon. As suggested by Moore et al. (2025), significant changes are needed in agriculture in order to increase soil carbon stocks.

## 8.2 Review of the study and further research

It is good to note that in reality, there is quite little data of the sites: from many sites, there is no data of the bulk density of the soil, or any soil samples from deeper in the soil. There are measurements from only two points of time, which is not enough to actually observe trends. In addition, measurements of carbon concentration from soil

samples are sensitive to measurement errors, and carbon concentration can vary a lot within a single field (Goidts et al., 2009). Therefore, too strong conclusions should not be drawn on the basis of these results.

On the other hand, the number of farms in the dataset is almost too big: collecting the farming data took a while, and in the end, some of the farms still lacked some farming data. Some farmers dropped out of the experiment. All this raises questions: Were all farmers motivated, and were carbon farming methods implemented until the end of the experiment? Was the data collected on the farming measures accurate? In this sense, the results from advanced sites might even be more reliable than from all sites. The data set is smaller, but farmers from advanced sites were known to be motivated.

In order to obtain more reliable results from a research approach such as the one presented in this thesis, the quality of the data should be given attention. Fewer farms could give better results, if the farming data was more reliable. Adding even just a third measurement of carbon concentration could allow the trend of the carbon stocks to be observed more reliably. Extending the experiment period could also improve the research, as changes in soil carbon occur over a long period of time and may not be visible in a short term. On the other hand, motivating farmers for a longer experiment could be difficult.

The modeling of the one-meter carbon stocks could also be improved. When calculating the prior distribution of the stock, the distribution of $c_0$ allows the topsoil carbon to get a wide range of results. This choice was made so that the prior does not steer the posterior distribution too much. Giving the prior a more realistic distribution, for example, one that does not allow negative values or values that are too high for a mineral soil, could still give better results.

The current results could also be analyzed further. Since there are farms from a wide area of Finland, it could be investigated whether different climatic conditions impacted these results. In addition to cover crops and nutrient-rich additions, the effect of multi-variety ley could be investigated, as it was the second most popular method in the experiment after cover crops. The comparison of treatment and control pairs could also be done individually, and differences between each pair could be further investigated using weather data and the recorded farming measures.

As using nutrient-rich additions showed promising results in decreasing the loss of carbon stocks, they should be further investigated. It could be investigated how much organic amendments needs to be added to prevent the carbon stocks of agricultural soils from declining, or even to increase them. This could be done similarly as here, using Yasso for simulating the carbon stocks. Complementing organic matter amendments with other carbon farming methods, such as cover crops, could also be investigated.

# 9 Summary

In this thesis, a carbon farming experiment conducted in 2019-2023 was studied. The experiment included treatment fields, where carbon farming practices were implemented, and alongside them control fields with conventional farming practices. Measurements of carbon concentration were taken from the fields before and after the experiment. The experiment was analyzed by quantifying the effect of carbon farming on photosynthesis and soil carbon inputs and stocks. Special attention was paid to farms that used cover crops or organic amendments as their carbon farming method.

The carbon stocks of the fields before and after the experiment were modeled assuming that the carbon content decreases exponentially with depth. Using Bayesian methods, the model was conditioned to carbon concentration samples taken from the fields. Modeling of the carbon stocks was evaluated using five-fold cross-validation, which showed adequate results. Carbon inputs during the experiment were calculated using plant productivity estimated based on satellite data, and records of yield harvested from fields and organic amendments applied to them.

The carbon stocks during years 1900-2025 were simulated using soil carbon model Yasso. The model was conditioned to observations of carbon stocks before and after the experiment and carbon inputs during the experiment, again using Bayesian methods.

From the simulated results, the years of the experiment were examined. The results showed that all of the fields were losing carbon, but slightly less in treatment fields than in control fields. According to our results, the treatment sites had less carbon than the control sites at the beginning of the experiment.

Using organic amendments had the biggest effect on decreasing the loss of carbon stocks and increasing the carbon inputs. Using cover crops showed a slight increase in the amount of photosynthesis. However, both in the treatment fields where cover crops were used and in their control fields, the amount of photosynthesis was low compared to other fields.

With the exception of organic amendments, the differences in the results between treatment and control were quite small. Additionally, soil samples were taken from each field at only two time points. Therefore, the results should be interpreted with caution. More accurate results could be obtained by increasing the amount of soil samples, and by examining the effect of different climatic conditions on these fields.

According to the results of this theses, organic amendments showed potential in decreasing the loss of carbon stocks of agricultural soils. In order to actually increase the carbon stocks of soils, more significant measures are needed.

# References


A. C. Abdullahi, C. Siwar, M. I. Shaharudin, and I. Anizan. Carbon sequestration in soils: the opportunities and challenges. *Carbon capture, utilization and sequestration*, 1:3–16, 2018.

A. Akujärvi, J. Heikkinen, T. Palosuo, and J. Liski. Carbon budget of Finnish croplands—effects of land use change from natural forest to cropland. *Geoderma Regional*, 2:1–8, 2014.

N. H. Batjes. Total carbon and nitrogen in the soils of the world. *European journal of soil science*, 65(1):10–21, 2014.

F. S. Chapin III, P. A. Matson, and H. A. Mooney. *Principles of Terrestrial Ecosystem Ecology*. Springer New York, New York, 2 edition, 2011.

C. Chenu, D. A. Angers, P. Barré, D. Derrien, D. Arrouays, and J. Balesdent. Increasing organic stocks in agricultural soils: Knowledge gaps and potential innovations. *Soil and Tillage Research*, 188:41–52, 2019.

T. E. Crews and B. E. Rumsey. What agriculture can learn from native ecosystems in building soil organic matter: a review. *Sustainability*, 9(4):578, 2017.

E. A. Davidson and I. L. Ackerman. Changes in soil carbon inventories following cultivation of previously untilled soils. *Biogeochemistry*, 20(3):161–193, 1993.

R. Derpsch, T. Friedrich, A. Kassam, and H. Li. Current status of adoption of no-till farming in the world and some of its main benefits. *International journal of agricultural and biological engineering*, 3(1):1–25, 2010.

European Commission. Cover and catch crops – general, 2022. URL https://wikis.ec.europa.eu/spaces/IMAP/pages/48169081/Cover+and+catch+crops_GENERAL.

European Union. Soil: The hidden part of the climate cycle, 2011. URL https://climate.ec.europa.eu/system/files/2016-11/soil_and_climate_en.pdf.

C. B. Field. Sharing the garden. *Science*, 294(5551):2490–2491, 2001.

Food and Agriculture Organization of the United Nations. What is soil organic carbon (SOC)?, 2025. URL https://www.fao.org/global-soil-partnership/areas-of-work/recsoil/what-is-soc/en/.

R. Gantlett, J. Bishop, H. E. Jones, and M. Lukac. Modern arable and diverse ley farming systems can increase soil organic matter faster than global targets. *Renewable Agriculture and Food Systems*, 39:e17, 2024.

A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472, 1992.

K. Georgiou, R. B. Jackson, O. Vindušková, R. Z. Abramoff, A. Ahlström, W. Feng, J. W. Harden, A. F. Pellegrini, H. W. Polley, J. L. Soong, W. J. Riley, and M. S. Torn. Global stocks and capacity of mineral-associated soil organic carbon. *Nature communications*, 13(1):3797, 2022.

E. Goidts, B. Van Wesemael, and M. Crucifix. Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales. *European Journal of Soil Science*, 60(5):723–739, 2009.

P. Günther, B. Garske, K. Heyl, and F. Ekardt. Carbon farming, overestimated negative emissions and the limits to emissions trading in land-use governance: the EU carbon removal certification proposal. *Environmental Sciences Europe*, 36(1):72, 2024.

R. Hay. Harvest index: a review of its use in plant breeding and crop physiology. *Annals of applied biology*, 126(1):197–216, 1995.

J. Heikkinen, E. Ketoja, V. Nuutinen, and K. Regina. Declining trend of carbon in Finnish cropland soils in 1974–2009. *Global change biology*, 19(5):1456–1469, 2013.

M. D. Hoffman, A. Gelman, et al. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623, 2014.

N. S. Jaikumar, S. S. Snapp, and T. D. Sharkey. Life history and resource acquisition: Photosynthetic traits in selected accessions of three perennial cereal species compared with annual wheat and rye. *American Journal of Botany*, 100 (12):2468–2477, 2013.

D. C. Johnson, R. Teague, S. Apfelbaum, R. Thompson, and P. Byck. Adaptive multi-paddock grazing management's influence on soil food web community structure for: increasing pasture forage production, soil organic carbon, and reducing soil respiration rates in southeastern usa ranches. *PeerJ*, 10:e13750, 2022.

K. Karhu, A. Wall, P. Vanhala, J. Liski, M. Esala, and K. Regina. Effects of afforestation and deforestation on boreal soil carbon stocks—comparison of measured c stocks with Yasso07 model results. *Geoderma*, 164(1-2):33–45, 2011.

J. P. Kaye and M. Quemada. Using cover crops to mitigate and adapt to climate change. a review. *Agronomy for sustainable development*, 37:1–17, 2017.

R. Lal. Soil carbon sequestration impacts on global climate change and food security. *science*, 304(5677):1623–1627, 2004.

J. Leifeld. Low-input farming: a way towards climate-friendly agriculture? *Carbon Management*, 4(1):31–41, 2013.

J. Liski, T. Palosuo, M. Peltoniemi, and R. Sievänen. Carbon and decomposition model Yasso for forest soils. *Ecological Modelling*, 189(1):168–182, 2005.

T. J. Mattila. Cover crops and soil loosening are key components for managing P and C stocks in agricultural soils. *Soil Use and Management*, 40(1):e12976, 2024.

T. J. Mattila, E. Hagelberg, S. Söderlund, and J. Joona. How farmers approach soil carbon sequestration? lessons learned from 105 carbon-farming plans. *Soil and Tillage Research*, 215:105204, 2022.

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, 2017.

S. B. Mirsky, W. S. Curran, D. A. Mortensen, M. R. Ryan, and D. L. Shumway. Control of cereal rye with a roller/crimper as influenced by cover crop phenology. *Agronomy Journal*, 101(6):1589–1596, 2009.

C. E. Moore, B. Blakely, T. L. Pederson, N. Gomez-Casanovas, C. D. Gibson, A. M. Knecht, G. Aslan-Sungur, E. H. DeLucia, E. A. Heaton, A. VanLoocke, T. Meyers, and C. J. Bernacchi. From depletion to restoration: Lessons from long-term monitoring of carbon gains and losses in cropping systems. *Global Change Biology*, 31(6):e70291, 2025.

Natural Resources Institute Finland. Evergreen revolution with cover crops, 2025. URL https://www.luke.fi/en/projects/ikivihrea.

J. K. Nyameasem, J. De Los Rios, C. Kluß, T. Reinsch, A. Poyda, F. Taube, and R. Loges. Incorporating leys in arable systems as a mitigation strategy to reduce soil organic carbon losses during land-use change. *Frontiers in Environmental Science*, 12:1399197, 2024.

T. Palosuo, J. Heikkinen, and K. Regina. Method for estimating soil carbon stock changes in Finnish mineral cropland and grassland soils. *Carbon Management*, 6 (5-6):207–220, 2015.

K. Paustian, E. Larson, A. Swan, E. Marx, J. Kent, and N. Zenes. Carbon farming : A working paper assessing the potential for soil C sequestration, 2017. URL https://acee.princeton.edu/wp-content/uploads/2017/07/Carbon-Framing-working-paper-23July2017.pdf.

K. Paustian, E. Larson, J. Kent, E. Marx, and A. Swan. Soil C sequestration as a biological negative emission strategy. *Frontiers in Climate*, 1:482133, 2019.

H. Susi, P. Thitz, M. Hagner, K. Raveala, J. Ekroos, and A.-L. Laine. Impacts of carbon farming practices on biodiversity at the farm scale, 2025. Preprint. EcoEvoRxiv.

M. Tuomi, T. Thum, H. J. Järvinen, S. Fronzek, B. Berg, M. E. Harmon, J. Trofymow, S. Sevanto, and J. Liski. Leaf litter decomposition—estimates of global variability based on Yasso07 model. *Ecological Modelling*, 220(23):3362–3371, 2009.

J. Vira, H. Vekuri, O. Nevalainen, M. Korkiakoski, T. Mattila, H. Aaltonen, M. Koskinen, A. Lohila, M. Pihlatie, and J. Liski. Improving agricultural carbon monitoring with sentinel-2 and eddy-covariance-based plant productivity estimates. *Carbon Management*, 16(1):2568042, 2025.

T. Viskari, J. Pusa, I. Fer, A. Repo, J. Vira, and J. Liski. Calibrating the soil organic carbon model Yasso20 with multiple datasets. *Geoscientific Model Development*, 15(4):1735–1752, 2022.

M. Wilson. The importance of parent material in soil classification: A review in a historical context. *CATENA*, 182:104131, 2019.