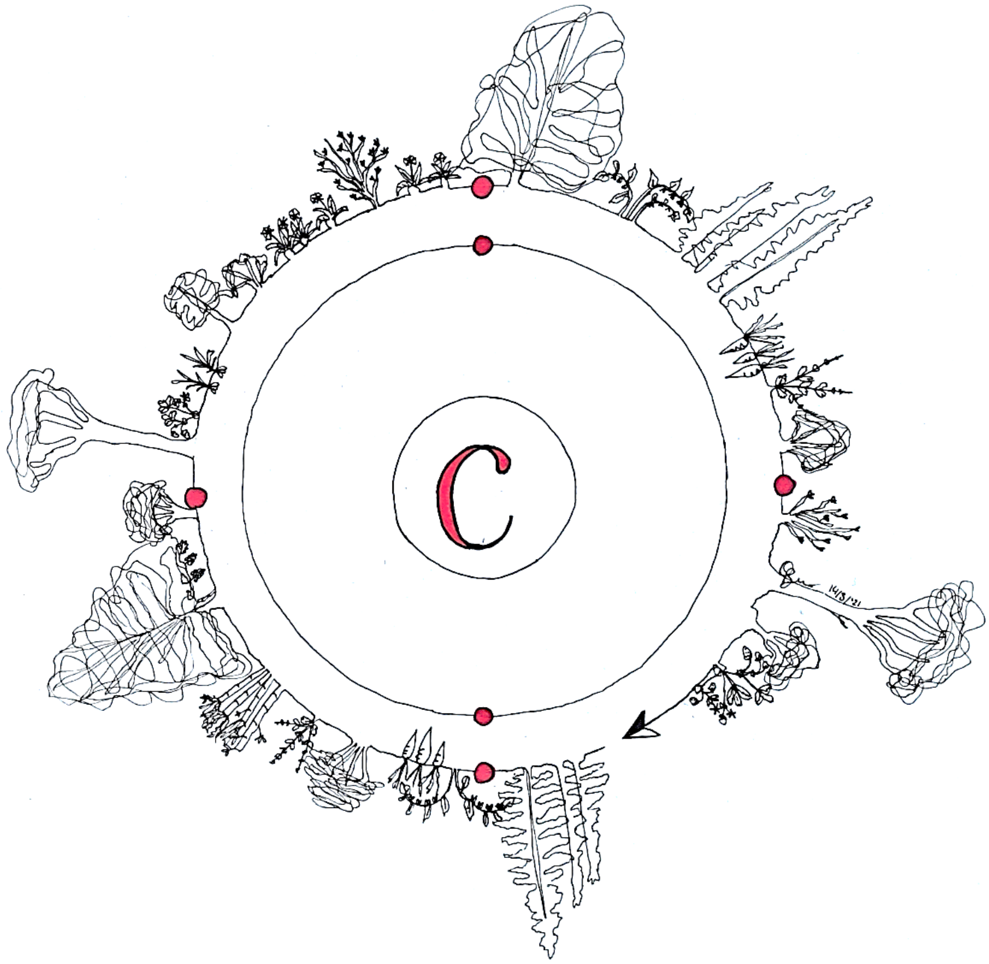
Assessment on the variation in aboveground carbon stock of Dutch food forests



K. Buinink
supervised by dr. Pita Verweij

MSc-programme Environmental Biology
Ecology and Natural Resource Management track
Utrecht University

Master Thesis
December 4, 2020

Utrecht University

Assessment on the variation in aboveground carbon stock of Dutch food forests

Major Research Project
Master program Environmental Biology
Ecology and Natural Resource Management track
Utrecht University, the Netherlands
Student number: 5719526

December 4, 2020

Supervisor	Bastiaan Rooduijn MSc
Examiner	dr. Pita Verweij
Second reviewer	dr. René Verburg

This study has been performed in collaboration with the
Dutch National Monitoring Program Food Forests
https://greendealvoedselbossen.nl/nationaal-monitoringsprogramma-voedselbossen/

The front page is illustrated by Quirine Brahma and symbolises the carbon atom as a common thread in the circular system of a food forest. All seven layers are represented in the illustration by the following species: *Pinus pinea, Juglans Regia, Torreya taxifolia, Malus domestica, Magnolia liliiflora, Diospyros kaki, Amelanchier ovalis, Ribes nigrum, Corylus avellana, Mentha spicata, Phyllostachys bissetii, Polygonatum multiflorum, Vitis vinifera, Fragaria x ananassa, Allium ursinum, Allium sativum* and *Daucus carota.*

Utrecht University

Nationaal Monitoringsprogramma Voedselbossen

Abstract

Food forests are an increasingly popular type of nature-inclusive agriculture. One of the ecosystem services provided by food forests is carbon storage. In this study, the variation in aboveground carbon stock of 21 food forests has been examined, as part of the Dutch National Monitoring Program Food Forests (NMPF). Furthermore, the effect of food forests on microclimate has been analysed and the variables to estimate aboveground carbon (AGC) stock have been tested for relevance. Mean aboveground carbon stock was significantly higher in former forests than in former arable lands and grasslands, with a difference of 62.8 and 64.8 Mg ha^{-1} respectively. Food forests with a *dispersed* structural composition had a significantly higher mean AGC stock than food forests with *alleys* (with a difference of 2.2 Mg ha^{-1}). No significant differences between three soil texture classes were found; nor a correlation between soil conditions and AGC stock. Carbon accumulation over food forest age was following a sigmoid curve, although the saturation point of this curve is uncertain. An analysis of AGC stock development over time within the same plots is necessary to confirm the sigmoid shape of the relation between AGC stock over time and to determine the exact curve. Food forests were both cooler (- 10.1 °C) and moister (+ 12.0%) than its environment. These differences were not increasing over food forest age or AGC stock. A generalised linear model including both maximum height of the tree layer (HT) and the total basal area (BA) of shrubs and trees explained over 99% of AGC stock. Both explanatory variables significantly explained over 80% of AGC stock variation individually (Adj. R^2 HT = 0.87, Adj. R^2 BA = 0.82). Although a model with only one variable was less accurate, maximum height and basal area might be applicable for low-threshold AGC stock estimations. The explanatory variable 'canopy closure' was less accurate. No fixed percentage of shrub carbon stock could be determined, and it therefore is useful to continue shrub measurements in further research on AGC stock of food forests. Further research is necessary to determine the which variables actually influence food forest aboveground carbon stock. This study could be used as baseline study for future research on carbon stock of food forests.

Layman summary

Over the last decades, much biodiversity has been lost. One of the reasons of this biodiversity decrease is the disappearance and degradation of habitat, mostly caused by agricultural land extension. Combining nature and agriculture is one possibility to restore biodiversity. In this way, land could be used for agriculture and nature can be restored at the same time. A food forest is such a type of nature-inclusive agriculture. Multiple species of trees, shrubs and herbs are grown together in a food forest, and since the use of pesticides and fertilizers have been left out, a natural system is able to develop.

The positive effect on biodiversity is not the only ecosystem service a food forest provides. Another one is the carbon storage of trees and shrubs growing in a food forest. Storage of carbon reduces the amount of carbon dioxide (CO_2) in the air and therefore combats global warming. In this project, carbon stocks of 21 Dutch food forests and what influences these carbon stocks have been assessed as part of the Dutch national monitoring program food forests (NMPF). The NMPF is established by the stakeholder cooperation Green Deal food forests in 2019. This is the first large-scale and standardised study on aboveground carbon (AGC) stock in Dutch food forests, which can form a basis for future research.

At least three sampling plots of 10x10m were randomly selected in every food forests. Species, height and stem diameter at breast height were documented for all trees within these plots. Species, height, crown diameter, number of stems and diameter of the three thickest stems were noted for all shrubs in these sampling plots. With these measurements, the total carbon stock in aboveground plant biomass was determined, so the roots of plants and soil organic carbon were not included. Next to this, the closure of the canopy, the temperature and humidity were measured, both inside and outside the food forest. In this way, the effect of the food forest on its microclimate has been examined.

After the fieldwork was performed, the results have been analysed. As hypothesised, the AGC stock was higher in relatively old food forests than in younger ones, with exception of former forests. These former forests had a significant initial AGC stock which did not increase over age. The relation between AGC stock and age followed a sigmoid curve for food forests on former grasslands and arable lands. This meant that carbon stock was first increasing exponentially, and saturated over time. Whether this sigmoid model actually describes the curve of AGC accumulation in food forests is uncertain. For example, the moment of saturation is questionable. The development of AGC stock within one sampling plot over time should be examined in future, to determine the accuracy of the sigmoid model.

As mentioned before, the relation between multiple variables and AGC stocks of food forests has been investigated as well. Since moisture and nutrient availability both positively influenced plant growth, and these conditions were correlated with soil texture, a significant effect of soil texture on AGC stock was expected. However, no relation between soil texture and AGC stock was found in this study. Neither a correlation between soil conditions (measured by the NMPF before) and aboveground carbon stock was found. Despite the absence of relations, one could not exclude the possibility of any effect of soil conditions on the accumulation of carbon stock. These relations might become visible when food forests are getting older, or when multiple annual measurements have been performed in each food forest. A significant relation was found between former land use and AGC stock. Former forests had a higher AGC stock than former arable lands and grasslands. The structural composition of the food forest did show a significant effect as well, as food forests where all crop species were planted mixed together (*dispersed*) had a higher AGC stock than food forests that were planted in rows (*alleys*). One could argue that this effect is caused by positive interactions between species, which are likely to occur more often in *dispersed* food forests. However, this hypothesis could not be confirmed yet.

This study did show an influence of food forests on their microclimate. Significant higher temperatures and lower humidity levels were found inside the food forest, compared to the outside conditions. These differences in temperature and humidity were hypothesized to increase over the age and AGC stock of food forests. However, this hypothesis could not be confirmed, possibly because of the lack of relatively old food forests and measurements errors at some of the sampling points. These measurements will be repeated in the coming years, after which the effect of food forests on their microclimate is to be analysed more thoroughly.

Finally, the accuracy of multiple carbon stock calculations has been examined. The methods used in this study were complex and time-consuming, but there might be an easier way to determine AGC stock. A model which included three variables (tree height, canopy closure and the sum of stem diameters) was found most accurate. Nonetheless, models including either maximum height or the sum of stem diameters were still quite precise. According to this, a fairly accurate estimate of AGC stock could potentially be made by measuring only one explanatory variable. This conclusion is yet to be confirmed in future, but it might be useful to distinguish two separate methodologies: a complex, time-consuming method to determine AGC stocks precisely versus a low-threshold alternative to estimate AGC stock quickly. There is no confirmation of whether it is useful to measure shrub carbon stocks. It might be possible to determine a fixed percentage of shrub carbon stock in future, which would make the relatively time-consuming measurements on shrubs redundant. This could not be done based on this study.

The findings of this study will be used as baseline for other studies part of the NMPF in the coming years. These results endorse the potential of food forests as sustainable, nature-inclusive type of agriculture, at least concerning carbon storage. This is, however, only the beginning of standardised, large-scale research on Dutch food forests and for many more to follow.

Nederlandse samenvatting

Als gevolg van het verlies aan leefgebied, is veel van de biodiversiteit verloren gegaan. Eén van de voornaamste oorzaken van dit verlies aan leefgebied is de uitbreiding van landbouwgebieden. Ook in Nederland. Eén van de oplossingen is het combineren van natuur en landbouw. Zo wordt de beschikbare ruimte nuttig besteed met landbouw en krijgt de natuur en de biodiversiteit kans om te herstellen. Een voedselbos is een vorm van zulke natuur-inclusieve landbouw. In een voedselbos groeien meerdere soorten gewassen door elkaar, zowel bomen, struiken als kruiden en wordt geen gebruik gemaakt van bestrijdingsmiddelen en bemesting. En dus krijgt de natuur de kans zich te ontplooien in een voedselbos: landbouw, maar dan inclusief natuur.

Naast het bevorderen van de biodiversiteit, is ook het opslaan van koolstof een groot voordeel van voedselbossen. Het opslaan van koolstof zorgt voor minder koolstofdioxide (CO_2) in de lucht en draagt dus bij aan het tegengaan van klimaatverandering. In dit project is onderzoek gedaan naar de koolstofvoorraden van 21 Nederlandse voedselbossen en gekeken naar waardoor koolstofvoorraden beïnvloed worden, allemaal onderdeel van het Nationaal Monitoringsprogramma Voedselbossen (NMVB) dat in 2019 door de Green Deal Voedselbossen is opgezet. Dit onderzoek is het eerste grootschalige en gestandaardiseerde onderzoek naar bovengrondse koolstof (BGK) opslag in Nederlandse voedselbossen, waar in de komende jaren op kan worden gebouwd.

In elk voedselbos zijn steekproefsgewijs plots van 10 bij 10 meter geselecteerd. In deze plots, minimaal 3 per voedselbos, zijn van alle bomen de soort, de hoogte, en de dikte van de stam op borsthoogte bepaald. Van alle struiken zijn de soort, de hoogte, de diameter, het aantal stammen en de dikte van de drie dikste stammen genoteerd. Dit samen leidde tot een schatting van de totale hoeveelheid opgeslagen koolstof in bovengrondse biomassa van bomen en struiken (alle bovengrondse delen van de planten; wortels zijn niet meegeteld). Ook is de geslotenheid van het kroondek, de temperatuur en luchtvochtigheid gemeten, zowel in het voedselbos als net daarbuiten, om te kijken wat het effect van het voedselbos op het microklimaat is.

Na het veldwerk zijn alle resultaten geanalyseerd. Zoals verwacht, was de koolstofvoorraad in oudere voedselbossen groter dan in de jongere voedselbossen. Uitzonderingen waren voedselbossen die ontwikkeld zijn in een voormalig bos: in deze voedselbossen was al veel BGK aanwezig en nam deze voorraad niet toe over de leeftijd. Voor alle andere voedselbossen gold: het verband tussen BGK-voorraad en leeftijd volgde een sigmoïde curve. Dit betekent dat koolstofvoorraad eerst exponentieel toeneemt en vervolgens verzadigt. Het moment waarop BGK-voorraad in voedselbossen verzadigt en of dit sigmoïde model ook daadwerkelijk klopt, kan nog niet gezegd worden. Om die vraag te beantwoorden moet worden gekeken naar hoe de BGK-voorraad in één en hetzelfde plot toeneemt over de komende jaren.

Zoals gezegd is ook gekeken naar het effect van verschillende variabelen op de BGK-opslag in voedselbossen. De verwachting was dat de bodemtextuur een duidelijk effect zou hebben op BGK-opslag. Want hoe vochtiger en voedselrijker de bodem, hoe sneller bomen en struiken kunnen groeien en hoe groter de BGK-voorraad is. Deze relatie tussen bodemtextuur en koolstofopslag is niet gevonden in dit onderzoek. Er is ook geen verband gevonden tussen bodemcondities, die zijn eerder al zijn gemeten door het NMVB, en BGK-opslag. Het is echter niet uit te sluiten dat deze bodemcondities toch een invloed hebben op hoe snel bomen groeien en hoe snel ze dus koolstof opslaan. Het voormalig landgebruik van voedselbossen had wel invloed op de BGK-opslag. Voedselbossen die gebouwd zijn in voormalige bossen hebben een grotere BGK-voorraad dan voedselbossen die gepland zijn op voormalige akkers en graslanden. De structuur van het voedselbos had ook effect op de koolstofopslag. Voedselbossen waar vele soorten door elkaar zijn aangeplant (*dispersed*) hadden gemiddeld een grotere BGK-voorraad dan voedselbossen waar alle gewassen in rijen zijn aangeplant (*alleys*). Dit zou kunnen komen omdat de verschillende boom- en struiksoorten een positieve invloed op elkaar hebben, maar het is te vroeg om dat te kunnen concluderen. Daarvoor is uitgebreider onderzoek nodig.

In dit onderzoek is verder gevonden dat voedselbossen het microklimaat beïnvloeden. In het voedselbos werden lagere temperaturen en hogere luchtvochtigheden gemeten dan 10 à 15 meter buiten het voedselbos. Ook was de verwachting dat deze verschillen in temperatuur en luchtvochtigheid toe zouden nemen naar mate de bomen en struiken in het voedselbos groter zouden zijn en er dus meer koolstof ligt opgeslagen. Dit verband is niet gevonden, maar het zou kunnen komen door het gebrek aan oudere voedselbossen of door meetfouten op sommige locaties. In de komende jaren wordt er steeds meer gemeten, en kunnen metingen van hetzelfde voedselbos uit verschillende jaren met elkaar worden vergeleken. Dan kan er een betere analyse gemaakt worden van of het effect van het voedselbos op het microklimaat verandert naarmate voedselbossen ouder worden.

Tot slot is er in dit onderzoek aandacht besteed aan de berekeningen van koolstofopslag. De huidige methoden zijn namelijk complex en kosten veel tijd. Wellicht is er een eenvoudigere manier om BGK-voorraad te bepalen. Het antwoord is tweeledig. Aan de ene kant is een model waarin zowel de hoogte van de bomen, de dichtheid van de kroonlaag als de dikte van alle stammen wordt meegenomen het meest precies. Maar aan de andere kant zijn zowel de hoogte van de hoogste boom in het plot als de som de dikten van alle stammen behoorlijk precies. Op deze manier zou relatief eenvoudig een redelijk accurate schatting kunnen worden gemaakt van de totale BGK-opslag in het voedselbos. In de komende jaren moet deze aanname nog worden bevestigd, maar het lijkt dus mogelijk dat er twee methodes komen: een ingewikkelde methode, die veel tijd kost, om daarmee heel precies de BGK-voorraad te bepalen en een eenvoudige methode, waarmee vrij snel een nauwkeurige schatting kan worden gemaakt van de BGK-voorraad. Ook over het wel of niet meten van struiken is nog geen uitsluitsel, ook daar is verder onderzoek voor nodig. Mogelijk kan er in de toekomst een vast percentage bepaald worden, waarmee het meten van alle struiken niet meer nodig is.

De methoden die gebruikt zijn in dit onderzoek en de resultaten die eruit voortgekomen zijn zullen de komende jaren gebruikt gaan worden voor verder onderzoek in Nederlandse voedselbossen als onderdeel van het NMVB. De eerste resultaten onderschrijven de kracht van voedselbossen als duurzame vorm van landbouw, in ieder geval wat betreft koolstofopslag. Maar dit is nog maar het begin en heel veel verder onderzoek zal volgen.

Acknowledgments

I want to start this report by thanking everyone who have helped me finish this major research project report. Starting with Bastiaan, for his endless inspiration, many sparring moments, and his faith. He gave me the opportunity to set up the research and perform all data collection myself, which made it entirely feel as my own research. I am proud that my commitment, results, and recommendations can be useful for the national monitoring study the coming years and thankful for the opportunity Bastiaan gave me to form the basis on carbon stock measurements in Dutch food forests. Second, I would like to thank my supervisor Pita. Although she was very busy, see conveyed her appreciation for my labour. Her pleasant and supportive way of coaching was encouraging me, and she knew to give a little push in the right direction at the appropriate moments.

But of course, my warmest thanks are to Fleur. Together, we managed to perform months of preparations, field work, data analyses and presentations, even though these were difficult and strange times. From day one, our cooperation has worked immaculate. You were there to help me if necessary, or rather to slow me down sometimes, but above all we have had a lot of fun together which made the fieldwork period much more enjoyable.

And then to all these people who have contributed to this project, to a greater or lesser extent. I am thinking of Maarten Schrama and Wouter van Eck for introducing me to their food forests and giving me the opportunity to evaluate my methodological approaches. I want to thank Wouter van Goor for contributing on the thought process of composing appropriate carbon stock measurements and AKVO foundation (Mert Blommenstijn in particular) for developing data collection software, which saved us valuable time in the field. I appreciated the help of Vincent de Leijster concerning the composition of a sigmoid model. And of course, I want to express gratitude to all food forest managers and volunteers for being a helping hand while we were collecting all data, their unbridled enthusiasm and the inspiring conversations. At last, I want to thank my lovely girlfriend Quirine and brother Fedde for advising me on the textual aspects of this report. To anyone who helped me, but which I could not mention here: I am grateful to you!

So, this is it. My major research project on the carbon stock of Dutch food forests is finished. I worked on it with great pleasure, and I hope that you will read it with just as much satisfaction. It is with great interest that I will follow the work and findings of the students who will continue the study on food forests as part of the national monitoring program. To all of them, good luck. We are just getting started!

Table of Contents

List of Figures and Tables

Figures

Appendix

Tables

Appendix

Abbreviations

adj. R^2	Value that described the proportion of variance predicted by the tested model, adjusted for the number of predicters in this particular model
AGC	Aboveground carbon (usually used as AGC stock)
AICc	Akaike Information Criterion
ANOVA	Analysis of variance, significance test for more than two categorical variables.
CDM	Clean Development Mechanism; methodology used in the Verified Carbon Standard
CEC	Cation exchange capacity (mmol / kg)
DBH	Diameter at breast height (130cm)
FF	Food forest
FLU	Former land use
glm	Generalised linear model
Mg ha^{-1}	Mega grams per hectare; unit of carbon stock
NMPF	National monitoring program food forests (NMVB in Dutch)
Ntot	Total amount of nitrogen in the soil (kg ha^{-1})
p	Probability of obtaining test results as extreme as actually observed; describes the outcome of the null hypothesis test
PCA	Principal component analysis
r	Correlation coefficient
REF	Reference forests
SOM	Soil organic matter (%)
VCS	Verified Carbon Standard

1. Introduction

The natural world as we all know is disappearing. Since the 1970s, over two third of the worlds' wilderness areas and 52% of its biodiversity (Planbureau voor de Leefomgeving, 2019; WWF, 2014; Brooks, 2002; IPBES, 2019) disappeared. Moreover, species extinction rates are up to one thousand times as high as they were for the last millions of years (De Vos *et al.*, 2015; Barnosky *et al.*, 2011). All these effects can almost completely be attributed to anthropogenic influences. Up to 50% of species loss is induced by habitat loss due to agricultural land extension (NEM, 2019; Geiger *et al.*, 2010; Planbureau voor de Leefomgeving, 2019). Rainforests are burned to provide soybeans for our cattle and palm oil for many of our daily products and diverse ecosystems are replaced by monoculture corn fields. The biodiversity crisis is not only facing the natural world, it is threatening humanity as well. The decrease in biodiversity has led to a decrease in essential ecosystem services (Díaz *et al.*, 2006; Cardinale *et al.*, 2012). For example, a decreased water storage potential of soils has increased the frecuency of floods (Wheater & Evans, 2009) and the decrease in insect abundance threatens crop pollination (Watt *et al.*, 1997; Collinge, 2000; Biesmeijer *et al.*, 2006). Moreover, global warming will be strengthened due to the decreased carbon sequestration capacity of our ecosystems (Malhi, 1999). In short, the consequences of biodiversity loss are problematic and biodiversity restoration is necessary to prevent the consequences from getting worse.

There are several ways in which biodiversity can be restored. For example, an increasing area of land and marine systems is classified as protected area (UN, 2018), cities become more insect friendly (IUCN, 2020; Dutch Ministry of Nature, 2020), and large areas of degraded land are reforested, including withered African desserts (Conservation International; Face the Future). However, the most destructive factor is the expansion of agricultural land. So, if we are determined to restore habitats, a change is needed regarding our agricultural activities. And since conventional agriculture is under pressure of changing environmental conditions as well (NCA, 2018), a switch to more sustainable farming techniques are undoubtedly imperative. A collaboration between human activities and nature can be the solution on both problems at the same time.

To conserve and restore nature, one could either separate of agriculture and protected nature reserves (land-sparing) or integrate both (land-sharing; Fischer *et al.,* 2014). Nature-inclusive agriculture might yield less than conventional agriculture in terms of food production, increases the required space for agriculture (Chappell & LaValle 2011; Foley *et al.,* 2005). However, the ecological value and biodiversity of these nature-inclusive agricultural lands are higher than on conventional agricultural farms (Hodgson *et al.,* 2010; Phalan *et al.,* 2011). Although less space could be entirely allocated to nature in this land-sharing principle (Green *et al.,* 2005), the implementation of nature in agricultural lands might be an effective way to conserve nature (Green *et al.,* 2005; Phalan *et al.,* 2011; Tscharntke *et al.,* 2012), even though critics prefer the idea of land-sparing (*e.g.* Kendal & Pimentel, 1994). Furthermore, one could argue that agricultural performance could be increased with the ecosystem services provided by nature. For example, microbial diversity can obviate the use of pesticides and soil water storage capacity can obviate irrigation of scarce fresh water (Cavoski *et al.*, 2011; Wang *et al.*, 2013). From another perspective, the harvest potential of this collaboration could make the investment in nature restoration more affordable. In this way, nature supports agriculture and vice versa.

A food forest is an example of nature-inclusive agriculture (Nero *et al.,* 2018; Park *et al.,* 2018; Riolo, 2019). In food forests, a large variety of tree, shrub and herb species are planted together in multiple strata (Green Deal Food Forests, 2019; Björklund, 2012; Nytofte & Henriksen, 2019). Compared to conventional agriculture, food forests are designed to function more like a natural system (Park *et al,* 2018; Nytofte & Henriksen, 2019). In a natural system, soil microbial community can develop better, increasing water storage potential and nutrient richness of the soil. This makes irrigation and fertilisation needless. Due to its high diversity, food forests attract more insects and birds than conventional agricultural systems, which provide natural pest control. Temperate food forests have originated in the early 1990s, making use of the principles of tropical home gardens and permaculture (Kehlenbeck, 2007; Ferguson and Lovell, 2013). Today, more and more commercial, and large-scale food forests are realised in Western-Europe, making food forests an increasingly popular sustainable agricultural concept (Voedselboskaart, 2020). Some of the pioneering food forest farmers claim to reach up to 10 times higher efficiency compared to conventional agriculture (pers. comm. Wouter van Eck). Both ecological and business advantages display the potential of food forests, although this is not scientifically proven yet.

In 2019, the Green Deal Food Forests, consisting of both governmental and non-governmental organisations, has established the Dutch National Monitoring Program Food Forests (NMPF; Green Deal Food Forests, 2019). Since the concept of food forests is relatively new in the Netherlands (or western Europe in its entirety), hardly any scientific research has been performed on this subject. The NMPF initially included twenty-one Dutch food forests, where standardised studies on a variety of ecological, economic, and social aspects were performed. The Dutch NMPF is the first large-scale research program on food forests.

One of the ecosystem services a forest provides is the capacity to mitigate climate change by permanently storing carbon (Brockerhoff *et al.*, 2017; Jose, 2009). And since food forests are expected to function as natural forests (Park *et al,* 2018; Nytofte & Henriksen, 2019), they are expected to have a large carbon storage potential as well. Because this research is part of the NMPF, a large variety of food forests is included. This study is the first in which aboveground carbon (AGC) stock was according to standardised protocols in Dutch food forests of different ages, with different former land uses, and on different soil textures. In essence, this is the first time that carbon stocks of food forests can be compared properly. As it is, this research can be seen as a baseline measurement on this subject, which can be built on in the future. The following research question is addressed in this study:

How does aboveground carbon stock in Dutch food forests develop over time and which ecological factors affect this process?

This study focusses on aboveground carbon stock only, no measurements or estimates on belowground carbon stock were made. The large heterogeneity between the included food forests enables the analysis of the effect of several variables on the AGC stock. One of those variables is the age of food forests. For most food forest, this study will be the first carbon stock assessment. Therefore, the current carbon stock could not be compared to previous years, and the development of carbon stock could not be determined within one food forest. AGC stock of young food forests will be compared with the carbon stock of older food forests, taking differences concerning other variables into account. Besides their age, food forests are selected based on their soil texture, former land use and structural composition (see *Section 3.1*). The effect of these variables on AGC stock of food forests will be evaluated, complemented with the interaction between these variables. To what extent the carbon stock food forests is comparable to the carbon stock of natural food forests will be discussed as well. Furthermore, the carbon stock measured in this study will be compared to the results of a study on food forest soils, performed on the same selection of food forests in 2019. Since forests influence their environment, an analysis of microclimate in food forests was made. Lastly, the suitability of measured properties (height, basal area, canopy closure and relative carbon stock of shrubs) to estimate aboveground carbon stock has been analysed.

For the sake of clarity, the broader scope of the research question was subdivided into several smaller-scale questions, namely:

1. *What is the relation between carbon stock and age of Dutch food forests?*
2. *Does carbon stock vary between food forests with a different soil texture, former land use or structural composition?*
3. *What is the relation between carbon stock of Dutch food forests and their microclimate?*
4. *What is the relation between carbon stock and soil conditions in Dutch food forests?*
5. *To what extent are basal area, maximum height, canopy closure and percentage of shrubs predicting aboveground carbon stock accurately?*
6. *Does the carbon stock of Dutch food forests differ from that of natural forests in the Netherlands?*

1.1. Hypotheses

The hypotheses concerning all six research questions mentioned above are briefly described in this paragraph. A more detailed explanation on these hypotheses can be found in the theory (*section 2).*

1.1.1. AGC stock over food forest age

As the trees and shrubs in a forest grow, the AGC stock will increase as well (Stephenson *et al.,* 2014). In general, the carbon accumulation of a forest follows biomass accumulation curve (Ciais *et al.,* 2008; Birch 1999), consisting of an exponential increase first, followed by a saturation (sigmoid). No scientific studies have been performed on the development of AGC stock in food forests before, but food forests are hypothesised to function as a forest and therefore carbon accumulation curves of food forests are expected to be comparable to these of natural forests. *It is hypothesised that the AGC stock of food forests increases over time, following a sigmoid model.*

1.1.2. Aboveground carbon stock per former land use, soil texture and structural composition

Plant growth rates, and therefore carbon accumulation rates (Stephenson *et al.,* 2014) are highly influenced by soil and environmental conditions (Basset *et al.,* 1964, Sullivan *et al.,* 2015). Clay soils are moister and nutrient richer than sand soils (Brown, 2007). Therefore, a larger AGC stock is expected in food forests on clay soils than in food forests on sand soils. Former land use is expected to influence AGC stock as well. Former forests have a larger initial carbon stock and are therefore expected to have a larger mean AGC stock than former arable lands and grasslands. Since microbial communities are more developed and diverse in grasslands than in arable lands (Schulte *et al.,* 2005; Girvan *et al.,* 2004), plant growth rates are expected to be higher on former grasslands than on former arable lands. Lastly, a *dispersed* food forest is expected to include more positive interspecific interactions than an *alleys*-design and therefore plant growth rates and AGC stock are expected to be higher in *dispersed* food forests than in food forests with *alleys* (Morin *et al.,* 2011; Palandrani, Battipaglia & Alberti, 2020). *In short, AGC stocks is expected to be influenced by all three categorical variables: soil texture, former land use and structural composition.*

1.1.3. Food forest microclimate

Temperature and humidity are both significantly influenced by forest vegetation (Lin & Lin, 2010). As food forests are expected to function like a forest (Park *et al,* 2018; Nytofte & Henriksen, 2019), the same effect is expected in food forests. *Significantly lower temperatures and higher humidity levels are expected inside food forests, compared to the outside conditions and these differences are expected to increase as aboveground carbon stock rises.*

1.1.4. Influence of soil variables on aboveground carbon stock

Plant growth is positively influenced by nutrient and water availability and an extensive and diverse microbial community (Bassett, 1964; Sullivan *et al.,* 2015; Craswell & Lefroy, 2001; Bot & Benites, 2005). Due to higher plant growth rates, a larger mean AGC stock is expected in presence of these favourable conditions. Furthermore, as a forest develops, the soil conditions change (Johnson & Wedin, 1997). The soil conditions of food forests with a large AGC stock are therefore expected to be significantly different from soil conditions of food forests with a small AGC stock. *It is hypothesised that AGC stock is significantly correlating with all soil conditions included in this research.*

1.1.5. Elements of aboveground carbon stock calculations

A wide range in AGC stock is possible concerning one of these variables, as trees and food forests are not expected to grow and develop similar. Therefore, calculations including all three variables are expected to be most accurate. Since allometric equations prefer the exclusion of height over the exclusion of basal area (Verschuyl *et al.,* 2018), basal area is expected to describe AGC stock more accurate than maximum height. Canopy closure is not even included in commonly used allometric equations (Verra, 2020; UNFCCC, 2013) and is therefore expected to explain variation in AGC stock worse than the other two variables. *In short, it is hypothesised that models including all three variables are explaining more variation in AGC stock and that basal area is the more accurate than maximum height and canopy closure.*

1.1.6. Food forests versus natural forests

In the climax stage of succession, the crown of a forest will almost completely be closed (Sigurdsson *et al.,* 2015). In general, food forests are expected to be designed without a complete closed canopy layer (Green Deal, 2019; Jennings, Brown & Sheil, 1999), in order to provide the herbaceous and shrub species with a sufficient light availability (Nytofte & Henriksen, 2019). Since most aboveground carbon is stored in the largest canopy trees (Stephenson *et al.,* 2014; Janssen *et al.,* 1999), *the carbon stock of food forests is not hypothesised to reach the carbon stock of natural forests.*

2. Theory

In this section, background information on several aspect was given. First, the Dutch national monitoring program food forests (NMPF) was briefly described because this study was part of this program. Next to this, some information on the Verified Carbon Standard (VCS) was given. The methodology on estimating aboveground carbon stock is mainly based on this standard. Lastly, a brief review on existing literature was given concerning all sub questions.

2.1. Dutch National Monitoring Program Food Forests

The Dutch National Monitoring Program was established by the Green Deal Food Forests in 2019. The Green Deal food forests is a cooperation between 25 stakeholders in order to join forces on the development of food forests in the Netherlands. These stakeholders included authorities (both ministries and provinces), multiple non-governmental organisations, the Dutch food forest foundation (Stichting Voedselbosbouw Nederland) and the research institutions NIOO (Dutch Institute for Ecology) and Wageningen University and Research (WUR). The goals of the Green Deal Food Forests include among other things the exchange of practical experiences, the organisation of renewed legislation and regulations concerning food forests and the bundling of research on food forests. The Green Deal Food Forests handles the following definition of a food forest: *"A productive ecosystem, designed by humans, after the example of a natural forest, including a high diversity on perennial and/or woody species producing food (fruits, seeds, leaves, stems, et cetera). At least the following things are present: 1. A layer of canopy trees. 2. At least three of the other vegetation layers (smaller trees, shrubs, herbs, ground cover plants, underground crops and climbing plants) 3. A significant and active soil life. A food forest has a size of at least 0.5 hectares in an environment with high ecological values. A minimum size of 20 hectares is handled in an ecological impoverished environment."* (Green Deal, 2020).

As part of this last goal, the NMPF was set up. This program facilitates large-scale and standardised research on all included Dutch food forests. At first, 18 food forests were included in this program, but this study already included 21 food forests and the goal of the NMPF is to grow and expand even further. As part of the NMPF, studies will be performed concerning multiple disciplines, including at least ecological, economic, and social assessments. These studies are intended to be linked together. For example, the study on soil conditions (De Groot, 2020) was used in this research.

As mentioned before, this study was performed as part of the NMPF. This is the first large-scale assessment on aboveground carbon stocks of food forests in the Netherlands. It is the intention to repeat the measurements on AGC stock in the following years, making use of the same methodologies.

2.2. Verified Carbon Standard

The Verified Carbon Standard (VCS) is a standard for the certification of carbon emission reductions. The VCS is controlled by the non-profit organisation Verra (Verra, 2020) since 2005. It is the most widely applied standard on carbon stock assessments (Verra, 2020). Other frequently used standards are the Gold Standard, which is mostly used by non-governmental organisations and Plan Vivo, which is mainly used in developing countries. These methodologies are less complex and expensive than the Verified Carbon Standard, but therefore also less accurate.

The VCS uses the methodologies composed by the United Nations Framework Convention on Climate Change (UNFCCC), which are called Clean Development Mechanism (CDM). The CDM methodologies are widely accepted to be one of the most accurate methodologies on carbon stock assessments. Multiple methodologies have been developed, for different kind of land types and assessments. The CDM methodology for both small-scale and large-scale afforestation and reforestation projects was determined to be most suitable for food forests and were therefore used in this research. However, the existing methodologies of the CDM have slightly been adjusted in order to enhance suitability for food forests.

Verra has set up over 20 validation and verification bodies across the world, to validate the composed methodologies and standards. Hardly any scientific and independent reviews on the general accuracy of the VCS are available, even though some studies did validate specific parts of the methodologies (Needelman *et al.,* 2018; Von Avenarius *et al.,* 2018; Sharma *et al.,* 2012). However, the VCS is frequently cited and used in a large variety of scientific papers. The VCS is therefore assumed to be accurate and the AGC stocks found in this research are well comparable to those of other studies that used the same methodology. It is however important to mention that no standard methodology is available for carbon stock assessments on food forests yet and one could argue whether an adjustment on the current methodologies is desirable.

2.3. AGC stock over food forest age

Aboveground carbon stock is correlated with forest age (Stephenson *et al.,* 2014). Older food forests are therefore expected to have a larger carbon stock than younger food forests. The curve of this carbon accumulation depends on tree growth rates (Stephenson *et al*., 2014), which are not constant over time (Birch, 1999; Brown, 1984). While growth rate of some tree species (for example increases as trees getting older, other species have a decreasing growth rate over age (Johnson & Abrams, 2009; Stahl & Urbance, 1990). Despite of these species-specific growth curves, a flattening of biomass expansion is found in at forest level (Ciais *et al.,* 2008; Birch, 1999). Because of this, the accumulation of carbon stock over the years is presumed to follow a sigmoid curve, as seen in temperate forests before (Dewar, 1990), mainly following the plant growth curve. The exact course of this curve, including the maximum sequestration rate and saturation level are dependent on planted species and climatic conditions and therefore generally hard to predict (Birch, 1999).

2.4. Aboveground carbon stock per former land use, soil texture and structural composition

Soil texture is hypothesised to influence carbon stock in food forests, since clay and sandy soils differ in amount of moisture, humus, and nutrients (Brown, 2007). Tree growth is positively influenced by a higher amount of moisture and nutrient availability (Basset *et al*., 1964; Sullivan *et al*., 2015), although another study nuances this (Besnard *et al.,* 2018). Food forests with higher tree growth rates built up carbon stocks more quickly. Therefore, food forests on clay soils are expected to have a larger carbon stock than food forests on sandy soils. Nutrient and moisture availability in loam soils are the intermediate between clay and sandy soils (Brown, 2007). Because of that, one could argue that the carbon stock on loam soils lays between clay and sandy soils as well. However, an excessively high moisture content could also negatively influence growth rates of trees (Predick, Gergel & Turner, 2009). These extreme conditions might presume in some of the selected food forests. In the end, the clay soils are hypothesised to have the largest carbon stock, followed by loam and sandy soils, respectively.

Former land use (FLU) is expected to influence carbon stock as well. Food forests cultured in a production forest or orchard will have a larger carbon stock than food forest cultured on arable and grasslands initially. Grasslands are more diverse and less exhausted than arable lands and will therefore provide more soil related ecosystem services, such as moisture retention (Schulte, *et al*., 2005). Moreover, grasslands have a larger and more diverse microbial community (Girvan *et al*., 2004). Tree growth and aboveground carbon sequestration are positively influenced by the amount of moisture and soil organic matter (McLauchlan, Hobbie & Post, 2006). Food forests on former grasslands are thus expected to have a larger carbon stock than food forests on former arable lands.

Two main structural compositions of food forests are distinguished: *dispersed*-designs, in which multiple crop species are planted mixed together and *alleys*-design, in which crops are planted in rows per species in order to maximise harvest usability. The effect of structural composition on the carbon stock of food forests is uncertain. In principle, plant species are able to encourage each other, which leads to a positive relation between plant diversity and carbon storage (Morin *et al.*, 2011; Palandrani, Battipaglia & Alberti, 2020). The more a food forest functions like a natural forest, the more positive interspecific interactions the system includes and the more carbon the food forest will store. Two structural compositions are distinguishable: *dispersed* and *alleys* (see *Section 3.1.1)*. In a dispersed food forest, multiple species are planted together and therefore looks more like a natural forest than a food forest with an alley-design. The alley-design is in turn more user-friendly concerning both harvest and management. Dispersed food forests are due to its design hypothesised to include more positive interspecific interactions and therefore store more carbon than food forests with alleys. However, there is no scientific evidence to substantiate this hypothesis yet.

2.5. Food forest microclimate

While a forest is growing, the effect of the ecosystem on the environmental conditions will increase as well (Lin & Lin, 2010). The presence of forests (or even individual trees) will lower air and soil temperature (Konarska *et al.*, 2016; Midrexler, Zhao & Running, 2011), and have an increasing effect on air humidity (Van Noordwijk *et al.*, 2014) in urban regions and in cities. Comparable effects are shown in agroforestry systems in Europe (Gosme *et al.*, 2016). However, natural grasslands have a cooling effect on temperature as well (Shen *et al.*, 2016), so what is the effect of planting a food forest on a grassland or arable land? During heatwaves in Europe, forests had a larger cooling effect in the long run than grasslands (Teuling *et al.*, 2010) and forests are assumed to be the major temperature buffering ecosystems in the world (Dalen, 2017). In this research, the temperature and humidity of the food forest are compared to the situation just outside the food forest. Although most food forests will be compared to grasslands in this way, a lower temperature and higher humidity caused by the trees, shrubs and/or dense vegetation is hypothesised. These differences are expected to increase with the development of the food forest (*i.e.* with carbon stock).

2.6. Influence of soil variables on aboveground carbon stock

Several significant correlations between aboveground carbon stock and soil conditions of food forests are expected. First, soil nutrient and water availability in food forests with FLU category *forest* significantly differ from soil conditions on former *arable lands* and *grasslands* (De Groot, 2020). Since food forests with FLU *forest* are expected to have a larger AGC stock, a significant correlation between carbon stock and nutrient and water content is hypothesised. Because a higher soil organic matter (SOM) content induces a high moisture and nutrient availability (Craswell & Lefroy, 2001; Bot & Benites, 2005), carbon stock is presumed to be correlated with soil organic matter as well. As mentioned before, food forest age is hypothesised to be correlated with carbon stock. Because of this, a correlation between food forest age and these soil conditions is expected as well, although former land use might influence this. In a principal component analysis (PCA), the components of the correlating variables are foreseen at the same axis. Concretely, the amount of soil organic matter and aboveground carbon stock are hypothesised to be depicted in the same direction in the PCA. The variable age is expected to point in the same direction as AGC stock and therefore in the same direction as the amount of SOM as well. Since former land use might strongly influence the correlation between age and SOM (De Groot, 2020), this effect might only be present when excluding food forests with FLU-category *forest*. Note that SOM is expected to be correlated with nutrient and water availability (Craswell & Lefroy, 2001; Bot & Benites, 2005), and that these arrows are hypothesised to be in the same direction as well.

2.7. Elements of aboveground carbon stock calculations

As food forests grow, the percentage of their canopy closure will increase (Sigurdsson *et al.*, 2005). With an increased canopy closure, the light availability for the understory will decrease (Jennings, Brown & Sheil, 1999). Since carbon stock will be correlated to tree growth, canopy closure in food forest is expected to be correlated to carbon stock. In a natural forest, crown canopy is entirely closed after 15 years on average (Sigurdsson *et al*, 2005). Therefore, the correlation between carbon stock and canopy closure only applies for the first stadium of the food forest, until complete closure is accomplished. Note that the closure rate of the canopy is not consistent for all food forests since food forests are expected to have a variation in growth rate (Bassett, 1964; Sullivan *et al.,* 2015) and the design (*i.e.* the proportion of canopy trees). However, it is assumable that canopy closure becomes relatively constant after a certain number of years (whether it is 15, or 20, or even more). Furthermore, some food forests might not reach a 100% canopy closure, as their design might prefer a relatively wide light availability (Green Deal, 2019; Jennings, Brown & Sheil, 1999). Again, this depends on the design and management of the specific food forest, making it hard to distinguish one specific curve of canopy closure over AGC stock. In short, canopy closure is hypothesised to be closely related to AGC stock for young food forests (excluding these with FLU-category *forest*), but it might be hard to use canopy closure as explanatory variable when forests become older.

Lastly, the relation between AGC stock and multiple properties of the food forests (canopy closure, height, basal area, and relative amount of carbon stored in shrubs) will be analysed. The relative presence of shrubs will decrease in a developing forest, based on the principles of forest succession (Urban & Shugart, 1992; Brooks *et al.*, 2012). In common carbon stock measurements, shrub carbon stock is excluded from the methodology (Verra; UNFCCC, 2013). Since a food forest principally contains multiple layers (Green Deal, 2019), the relative amount of carbon stored in shrubs is assumed to be higher than in natural forests. Although the relative carbon stock of shrubs will reduce as total carbon stock increases (Urban & Shugart, 1992), it is hypothesised to remain a significant fraction of total carbon stock and will therefore be necessary part of carbon stock measurements in food forests. Canopy closure, maximum height and basal area are all expected to largely explain variation in AGC stock. However, none of these variables is expected to be appropriate to predict AGC stock accurately. Therefore, a combination of these variables is hypothesised to be necessary to determine AGC stock properly. Though it is interesting to determine which variable is most precisely describing AGC stock and to what extent this variable can be used for quick, rough estimates of carbon stocks of food forests in future.

2.8. Food forests versus natural forests

On average, a Dutch forest has an AGC stock of 59 Mg C ha^{-1} (Nabuurs & Mohren, 1993). Food forests are not expected to approach the aboveground carbon stock of more natural forests (Schafer, Lysák & Henriksen, 2019. Food forests in general include both shrub and herbaceous crops in their design, requiring a certain light availability (Green Deal, 2019; Nytofte & Henriksen, 2019). The density of canopy trees is therefore expected to be less than in a more natural forest. Since most carbon is stored in the canopy trees (Stephenson *et al.,* 2014; Janssens *et al.,* 1999), the maximum carbon stock of food forests is expected to be lower than the maximum carbon stock of natural forests (due to this design; Lehmann *et al.*, 2019; Schafer *et al.*, 2019). Note that there are many possible food forest designs and that some food forests might develop a completely closed canopy, accepting a low light availability at forest floor. The oldest food forest in the Netherlands is 22 years old, which is significantly lower than the age of forests in the study of Nabuurs & Mohren (1993), that had a mean age of 50 years. The maximum carbon stock in food forests in this research is therefore hypothesised to be more similar to the food forest in Devon, UK, which has an age of 23 years and an aboveground carbon stock of 34 Mg C ha^{-1} (Lehmann *et al.,* 2019). The aboveground carbon stock of tropical home gardens and agroforestry systems is not expected to be reached in food forests in the Netherlands, since aboveground carbon stocks in temperate forests do not equal carbon stocks in tropic forests in general (Malhi *et al*., 1999). It is important to mention that the proportion of carbon stored in AGC carbon stock is significantly higher in tropic forests (Malhi *et al.,* 1999; Trumbore, 1993; Rooduijn *et al.,* unpublished data). If the AGC stock of temperate (food) forests is smaller than its tropic equivalent, this does not automatically mean that the ratio of total stored carbon (aboveground and belowground) is comparable.

3. Methods

As mentioned before, this study is part of the Dutch national monitoring program food forests (NMPF). Since this is the first large-scale study on carbon stock in Dutch food forests, an appropriate method had to be developed. These methods are mainly based on similar existing research, adapted for food forests in the Netherlands based on the expertise of specialists. At all times. The feasibility of the methods as part of the NMVB must always be considered. The used methods are therefore a compromise between the most accurate and the most accessible way of estimating the carbon stock of a food forest. However maximum accuracy under given circumstances is pursued. In this section of the report, the following methodological aspects are described:

1. Plot selection
2. Data collection
3. Carbon stock calculations
4. Data analyses

3.1. Study site and plot selection

This research includes 21 food forests across the Netherlands (see *Figure 1*). This selection has been made by the Green Deal Food Forests, among other things based on age, location, soil texture and former land use. An overview of included food forests, including these variables can be found in *Appendix 1;* see *Figure 2* & *4* for impressions of a variety of included food forests. Plots of 10x10m were used for the carbon stock measurements since this plot size is most adopted by ecologists for

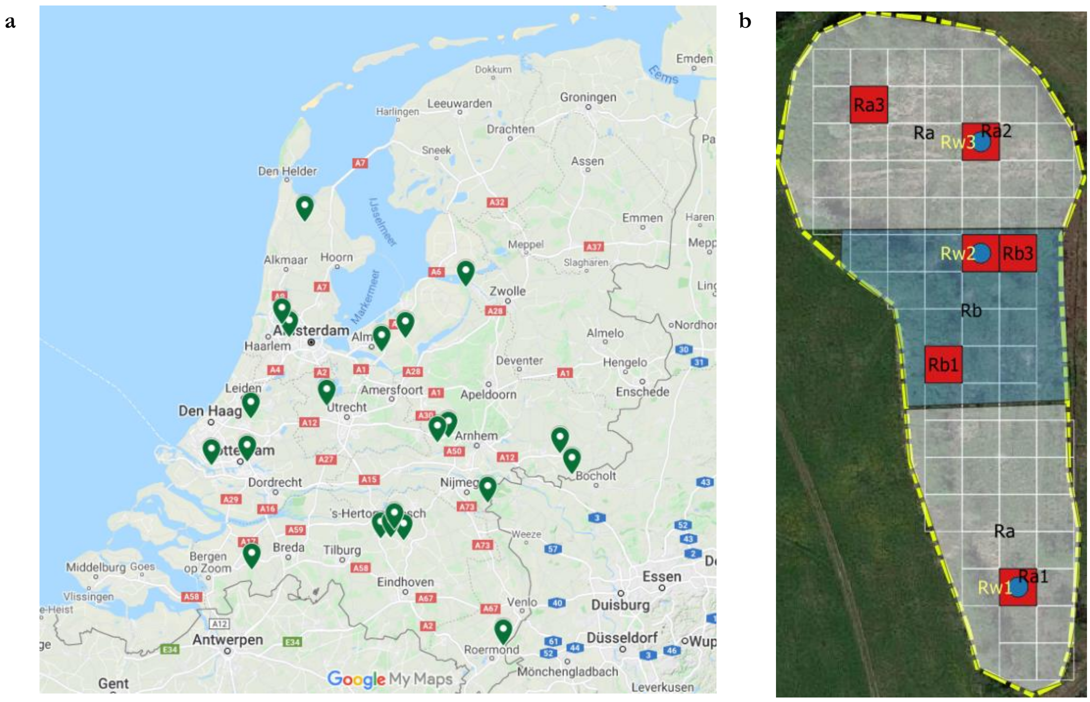


Figure 1: Sampling locations of the 21 food forests. a) Location of the 21 food forests in the Netherlands. Food forests are located in six provinces: Noord-Holland, Zuid-Holland, Flevoland, Utrecht, Gelderland, Noord-Brabant and Limburg. *Exported from Google My Maps.* **b)** Grid and selected sampling plots for food forest Roggebotstaete. Red squares represent 10x10m sampling plots. Blue and grey areas represent two zones based on soil texture. Yellow codes and blue dots refer to soil measurement locations of the NMPF. *Exported from QGIS.*

many years (Bourdeau, 1953). To determine the location of the measurements, all food forests were divided into a grid of plots of 10x10 metre in QGIS (*see Figure 1;* QGIS, 2020). Plots that were partially located outside of the food forest border are deleted. Furthermore, plots (partially) located in a pond, at a paved path or at buildings are excluded from the grid. Using a random selection tool in QGIS, the requested number of plots were selected randomly out of all available plots in the grid (see *Appendix 2* for the random selection protocol).

The number of plots depends on two things. First, the size of the food forest determines the number of measuring plots. In the most desirable scenario, the number of measured plots per hectare is equal for all food forests. The appropriate number of plots is determined at three plots ha^{-1} (CDM, 2010). Of course, a greater sampling effort would increase the accuracy of the measurements, however this is not accessible in the national monitoring program. Therefore, the minimum number of three and the maximum number of six measurement plots were allocated to the food forests. Since the size of most food forests included in this research is between one and two hectares, the ideal number of plots per hectare has been used in most of the sampling locations. Second, some of the food forests were divided into zones (see *section 3.4.* which food forests were split up). In this scenario, the number of measurements plots will be three plots per zone, in dependent of the zone size.

Figure 2: Photographs of four food forests with a variety in age. a) Food forest Schijndel Hardekamp located in Schijndel, Noord-Brabant with an age of 1 year. **b)** Food forest Mijn Stadstuin, located in Amsterdam, Noord-Holland, with an age of 5 years. **c)** Food forest Ketelbroek, located in Groesbeek – De Horst, Gelderland, with an age of 12 years. **d)** Food forest Sualmana, located in Swalmen, Limburg, with an age of 22 years.

3.1.1. Determination of food forest zones

Food forests are generally characterised by a large heterogeneity (Jacke & Toensmeier, 2005). By using random selection and a minimum number of three plots per food forest, the measured carbon stock was assumed to be representative for the carbon stock of the entire food forest, despite of the heterogeneity. However, in some scenario's, the food forests included in this research were divided into multiple, clearly distinguishable zones. One could argue that every food forest is divisible into zones, and in essence this argument is true. Therefore, several scenario's in which it might still be useful to split up the food forest were determined. Using these zones, differences within one single food forest could be analysed. Ideally, two of these zones are mostly under the same conditions, except of one or two of these conditions. In this way, it might be easier to analyse the effect of single variables on the carbon stock of food forests.

Food forests were divided in zones based on one of the following five variables. The exact protocol concerning zone selection can be found in *Appendix 2.5*.

- Age of the food forest
- Former land use
- Soil texture
- Elevation
- Structural composition

3.1.2. Reference plots

Although the main objective of this research is to describe the differences in carbon sequestration between food forests, the carbon stock in food forests was compared to the carbon stock of natural forests as well. Although it was impossible to determine whether a forest is a natural forest, it was interesting to compare its carbon stock with the carbon stock of food forests, especially those that were converted from a production forest or arboretum (i.e. with former land use 'forest'). Therefore, three measuring plots were selected in forests, located next to (or close to) a food forest. Since it was hypothesised that the carbon stock in forests depends on soil texture (Bassett *et al.,* 1964; Sullivan *et al.,* 2015; Brown, 2007), one reference forest was selected for every soil texture class each (three in total). To prevent the sampling effort for being too large, only one plot was measured in all reference forests each. Since the heterogeneity in 'natural forests' is lower than in food forests (Asase & Teteh, 2010; Brüning *et al.,* 2018), the observations were assumed to be quite representative, even by using one measuring plot only.

As in the food forests itself, the plots were selected randomly in the reference forests as well. Due to the absence of detailed maps (which were necessary for grid selection in QGIS), the methods for random selection in reference forests were different. From one corner of the forest (by default the south-eastern corner), one started walking northwest for one minute. After one minute, a stopwatch was randomly stopped, and the first two decimals were noted. These numbers had to be walked north and west respectively, to reach the north-eastern corner of the measuring plot. This method is commonly known as a convenient way to select a sampling plot randomly and to prevent for subjective decisions (Maltamo *et al.*, 2011; Ruotsalainen *et al.,* 2019; Wilson, 2005). The methods used in the reference plots were the same as in the food forest plots.

3.2. Data collection

3.2.1. Measurements on trees and shrubs

In the 100 square metres plot, all trees and shrubs are measured. Since the results had to be comparable to other studies, the methods included in the verified carbon standard (VCS) and its measurements (CDM) are used as much as possible (Verra; UNFCCC, 2013; UNFCCC, 2010). However, these standards are designed for studies in production forests and reforestation projects. A food forest design generally contains a larger number of shrubs than a reforestation or production forest (which are mostly focussed on maximum carbon storage). Therefore, shrubs were added to the methods. Furthermore, the VCS only includes trees from (at least) five centimetres of diameter at breast height (1.30m above the soil). Since most of the food forests in this research were only several years old, trees were included from five millimetres of stem diameter. In this way, the development of carbon stock in the first years of the food forest was estimated more accurately.

All measurements on aboveground carbon stock are visualised in *Figure 3*. Following the CDM methodology, all trees, and shrubs of which the stem was over half located within the plot were included (doubtful cases were excluded; UNFCCC, 2013). As shown there, the measurements are not the same for trees as for shrubs. The calculations for carbon stock in shrubs required other components than those of trees. As both calculations were different, it was important to standardise the way of calculating the carbon stock of each species. Prior to the fieldwork, a list of species has been composed to determine whether that species is a tree or a shrub, based on their dominant growth form (see *Table A3*). Even if a so-called shrub species has grown like a tree in a specific situation, that individual has been measured as a shrub. In this way, the carbon stock has been estimated equally in all food forests.

Carbon stock in tree calculations are mainly based on the diameter of the tree stem at breast height (1.30m). The diameter was measured using a diameter tape and perpendicular to the axis of the tree stem (note that this was not always horizontal). Since young trees were included in this research as well, the stem could be branched-off at a height lower than 1.30 metres. If there was still one upward growing stem, this stem was measured, despite of the branching of the stem. If there was no central stem at 130 centimetres, the diameter of the stem was measured at 60 centimetres or lower alternatively. If a tree was growing out of multiples stems from the ground, these trees were measured as separate trees. The height of the tree was determined using an inclinometer. The inclinometer has given an angle, which is converted to the height of the tree, using the following formula:

$$H = (\tan\alpha \cdot D + L) * 100 \qquad \textit{(Equation 1)}$$

where:

H = tree height in centimetres
α = angle given by the inclinometer
D = distance to the tree in metres (which was fixed at 10m)
L = eye level of the observer in metres (which was fixed at 1.65m)

It was important that there were no differences in altitude between the measuring point and the tree. When trees were smaller than 200 centimetres, it was easier and more accurate to measure tree height directly, instead of using an inclinometer.

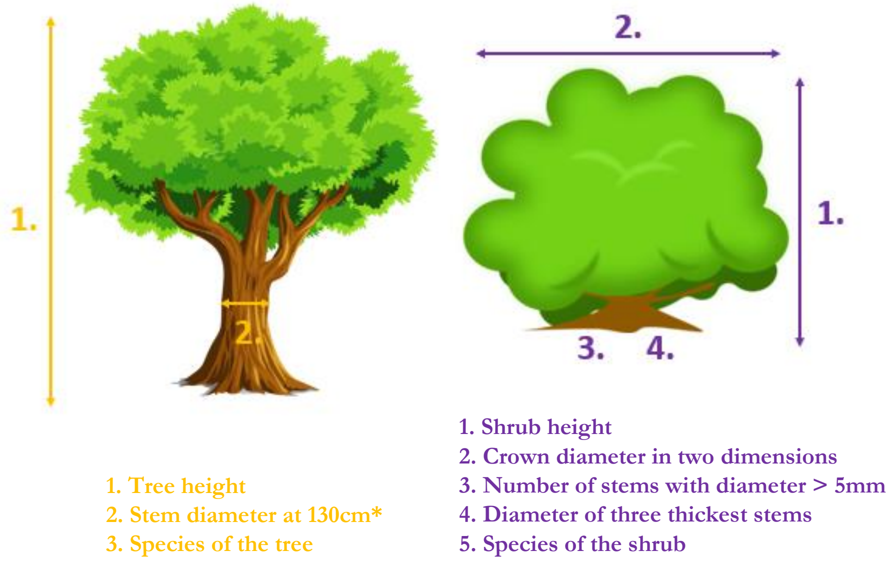


Figure 3: Scheme of all measurements that have been taken on all trees and shrubs within the 10x10m plots.* When tree stem was branched off at 130 centimetres, stem diameter was measured at 60 centimetres, or just below the first branch point. Tree and shrub species were determined using Heukels' Flora van Nederland (Van der Meijden, 2005).

The measurements for shrubs were more complex than those for trees, since allometric equations of shrub biomass require more variables (Verschuyl *et al.,* 2018). First, shrubs carbon stock calculations are dependent of their crown cover. Since the crown of shrubs could grow ellipse-shaped, the shrub crown was measured over two perpendicular axes. Furthermore, it was also important to document the number of stems. If shrubs had more than three stems from the ground, only the three thickest stems were documented (to prevent the methods for being too time-consuming) and the basal area of the shrubs were calculated later, based on the available information. The diameter of the stems was preferably measured at a fixed height. However, due to the large variety of branching forms of shrubs, this was hardly applicable. There was tried to measure stem diameter in the range of 10-30cm, but extremely low ramifications could make this impracticable. The height of shrubs was mostly measured directly since most shrubs are less than 200 centimetres in length. Despite of this, it was recommended to use an inclinometer for the larger shrubs as well.

3.2.2. Microclimate measurements

In addition to the aboveground caron stock measurements, the microclimate of the food forests was observed. For this microclimate analysis, the temperature, humidity, and light availability were documented. Since light availability is directly correlated to the closure of the canopy layer (Jennings, Bron & Sheil, 1999), the amount of available light was represented by the percentage closure of the canopy. After the application Canopy Capture (Patel, 2018) was tested during the preparatory weeks, this application was assumed to be precise enough to measure canopy closure. Although this made the measurements more accessible, this assumption should be tested in future research. The application was opened, and the mobile device was held horizontal at 1.5 metres, after which a picture was taken. Although the application came up with a percentage, this should not be adopted immediately. Especially in changeable weather conditions, this percentage might be incorrect. Therefore, the picture was always checked, and the correctness of the application outcome was verified. Since the closure of the canopy layer could vary significantly within several metres, the canopy closure was measured at three spots in each plot, namely the northwest corner, the southwest corner, and the centre of the measuring plot. These three percentages were averaged to determine the canopy closure of this plot.

The temperature and humidity of the air were measured using a thermo hygrometer (Lascar Electronics). This device automatically registered both temperature and humidity each ten minutes. Since the effect of the food forest vegetation on these conditions was more interesting than the absolute temperature and humidity itself, a control measurement outside the food forest was taken as well. This device was placed outside of the food forest, within 20 metres of the food forest border. Furthermore, this device was not placed on paved paths or roads and not overshadowed by any trees. The other device was used for the measurements inside the food forest. This device was placed in the centre of the measuring plot. Because of the large heterogeneity within food forests, it was necessary to translocate the thermo hygrometer along all measuring plots. The exact moment of picking up the thermo hygrometer were documented, since the device is registering constantly. Note that it was important to minimalise the time of carrying the device, to prevent the data from being influenced. The data of moments in which the device was not located in the vegetation, for example during moving moments, were manually removed from the database after the measurements were completed.

For two reasons, the closure of the canopy was measured as well. First, canopy closure is indicating light availability for herbaceous vegetation. And second, canopy closure is hypothesised to be related to aboveground carbon stock. In each sampling plot, canopy closure will be measured three times: the first in the plot centre, the second one metre from the north western corner towards the plot centre and the third one metre from the south eastern corner towards the centre of the plot. Canopy closure will always be measured at 1.50m. The application Canopy Capture (Patel, 2018) was determined to be as accurate as a densitometer during test measurements and therefore appropriate to be used in this research.

3.3. Carbon stock calculations

After measuring all required variables, the carbon stock of aboveground biomass was calculated. To do so, allometric equations were used. Furthermore, it was necessary to determine the wood density of all species observed in the food forests. The ICRAF Database combines many databases and studies on wood density and is therefore the most complete available database (World Agroforestry, 2020). By default, this database was used for all wood densities used in this research. If species specific wood density was unknown, the mean wood density of the genus where this species belongs to was used instead. If any wood density of that genus was missing as well, the mean density of the plant family was used. In the most unfavourable scenario of missing an entire family in the database, the mean wood density of all species in the database was used.

3.3.1. Tree allometric equations

To translate the tree dimensions (i.e. height and stem diameter) to an amount of carbon stored in aboveground biomass, the following allometric equations were used:

$$C_{\mathrm{TOT}} = \frac{1}{10^6} \cdot \frac{\sum C_{\mathrm{TREE,j}}}{A} \qquad \textit{(Equation 2)}$$

$$C_{\mathrm{TREE,j}} = \frac{44}{12} \cdot B_{\mathrm{TREE,j}} \cdot cf \qquad \textit{(Equation 3)}$$

$$B_{\mathrm{TREE}} = V \cdot Dw_{\mathrm{j}} \cdot BEF \qquad \textit{(Equation 4)}$$

where:

C_{TOT}	= amount of stored CO_2 in aboveground biomass (t CO_2 ha^{-1})
A	= sample area size(m^2)
$C_{TREE,j}$	= amount of stored CO_2 in aboveground biomass of species j (g)
$B_{TREE,j}$	= aboveground tree biomass of species j (g)
cf	= carbon fraction of tree biomass
V	= volume of tree stem (cm3)
Dw_j	= species specific wood density of species j (g cm^{-3})
BEF	= biomass expansion factor

The volume of the stem was calculated using the formula of the volume of a cylinder ($\pi \cdot r^2 \cdot h$), in which h is the tree height in cm and r is half the stem diameter in cm. Genus or family mean wood densities were used in absence of species specific wood densities (as described above). The default biomass expansion factor of 1.15 and the default carbon fraction of 0.47, as used in the CDM methodology (UNFCCC, 2013), were used for all species in this research. To obtain the amount of stored carbon dioxide, the stored carbon was multiplied with $\frac{44}{12}$, which is the relative molecular weight of carbon dioxide to carbon. In the end, the amount of stored carbon dioxide in all trees was added together and divided by the sampled area size (which depended on the number of measuring plots). The amount of carbon dioxide was divided by 10^6 to convert the unit of stored CO_2 from g ha^{-1} to the commonly used t ha^{-1} (or Mg ha^{-1}).

3.3.2. Shrub allometric equations

Where the allometric equations of tree biomass and carbon storage are widely used and standardised in the Verified Carbon Standard, the carbon stock in shrub biomass is less used in carbon stock estimated. To calculate the aboveground biomass in shrubs, allometric equations based on a research at the Oregon Coast Range are used (Verschuyl *et al.,* 2018). In this research, carbon stock in shrubs is determined using destructive methods and predictive models based on the dimensions of the shrubs are made. The climate of the Oregon Coast Range is only slightly different from the climate in the Netherlands (NCEP, KNMI) and the species included in their research are largely corresponding or comparable to the species found in food forests. Furthermore, Verschuyl *et al.* conclude that these allometric equations are applicable in other, comparable fields. Therefore, these allometric equations are assumed to be most suitable available equations.

As we do, Verschuyl *et al.* included all stems from 5mm in diameter size. However, stem diameters were measured at a fixed height of 15cm, while the measuring height was varying in this research, mainly due to differences in growth form of shrubs. The equations included three variables of shrub dimensions: shrub height (HT), crown area (CA) and basal area (BA). The height of shrubs was measured directly in food forests. To determine the crown area, the measured crown diameter in both directions was implemented in an ellipse formula:

$$CA = \pi \left(\frac{CDx}{2}\right)\left(\frac{CDy}{2}\right) \qquad \textit{(Equation 5)}$$

where:

CA	= crown area in m^2
CDx	= crown diameter in dimension x in m
CDy	= crown diameter in dimension y in m

Shrub basal area was not directly measured in food forests, due to limited time and the objective to make the methods accessible. However, the predictive capacity of the models that included BA were higher than those that were only based on CA and HT (R^2 of 0.93 and 0.63-0.68 respectively). Since the three thickest stems were measured, the BA was easily calculated for shrubs with three or less stems by summing up the area of each stem ($V\text{STEM} = \pi \cdot r^2$), in which r is half the stem diameter. For shrubs with more than three stems, a linear growth model in Microsoft Excel was used to estimate the BA (Microsoft Corporation, 2018). This model assumed a linear decrease of the stem diameter size. This decrease was calculated within the three thickest stems and using linear extrapolation, the diameter of the other stems was estimated. Since we have only counted stems with a diameter of more than 5mm, all estimated basal stem areas of less than 1.989 cm^3, which corresponded to a diameter of 5mm, were manually set to a basal area of 19.635. All stem basal areas were summed up to obtain the estimated basal area of those shrubs.

For some shrubs (less than 9% of all individuals), the number of stems and/or the diameter of the three thickest stems was undeniable. For these shrubs, the allometric equation without BA was used (Verschuyl *et al.,* 2018), although the predictive capacity of this model was less than the model for the other shrubs. The following equations were determined to be the best predictive (Verschuyl *et al.,* 2018), with a corresponding R^2-value of 0.93 and 0.63-0.68 respectively:

$$B\text{SHRUB} = 1.1888 + 0.7292\ \text{BA} + 0.30406\ \text{CA} + 0.4185\ \text{HT} \qquad \textit{(Equation 6)}$$

$$B\text{SHRUB} = 6.2855 + 1.1717\ \text{CA} \qquad \textit{(Equation 7)}$$

where:

BSHRUB	= Aboveground shrub biomass in g
BA	= Basal area in mm^2
CA	= Crown area in m^2
HT	= Shrub height in m

The equation without basal area as predictive variable, does not include shrub height as well, since the R^2 of an equation with shrub height was lower than the equation shown above. The shrub biomass was converted into shrub carbon stock at the same way as this was done for trees, using *Eq.2* and *Eq.3*.

3.4. Categories of food forests conditions

The main question of this research is how carbon stock accumulates over time in food forests. Therefore, it is necessary to determine the age of all food forests included in this research. Furthermore, the influence of both categorical variables soil texture and former land use on the carbon accumulation were investigated. In this paragraph, the categorisation of food forests based on their environmental conditions will be described.

3.4.1. Former land use

The selection of food forests included in the national monitoring program (and therefore in this research) was amount other things based on their former land use. In essence, what is the history of the land management in the years prior to the conversion into a food forest. To make it more understandable, the focus was on the last several years and further history was ignored. The former land use was expected to have a significant influence on the initial soil conditions, and therefore on the growth of food forest vegetation within the first years (Schulte, *et al.,* 2005; Girvan *et al.,* 2004; McLauchlan, Hobbie & Post, 2006).

As shown in *Table 1*, a large variety of former land use types was included in the dataset (11 types in total). Since this is hard to use in statistical analysis and there were only slightly differences between some of them, all these former land uses were united in three categories: *grassland*, *arable*

a b c

Figure 4: Photographs of three food forests with a variety in former land use. All three have an age of 2 years. **a)** Food forest Groengenoten, located in Rucphen, Noord-Brabant; a former grassland. **b)** Food forest Schijndel Boschweg, located in Schijndel, Noord-Brabant; a former arable land. **c)** Food forest De Stomp, located in Westendorp, Gelderland; a former production forest.

land, and *forest*. The FLU category *forest* was characterised by the presence of trees and/or shrubs even before the food forest was planted. The difference between grassland and arable land was a bit more complex. The FLU category *arable land* was defined by lands with an intensive land management for growing crops, with or without the usage of pesticides and fertilizers. Grasslands were less intensively used and were mainly characterised by the grazing of livestock. Note that the number of food forests in the second column is higher than the number of food forests included in this research (23 and 21 respectively), since some of the food forests had multiple former land uses.

Table 1: Categorisation of 11 former land uses into three former land use categories: Grassland, arable land, and forest.

Former land use category	*Number of food forests*	*Included former land uses*
Grassland	7 (30%)	Production grassland, natural grassland, horse field, sheep pasture, recreation field
Arable land	10 (44%)	Arable land (unspecified), corn field, grain field
Forest	6 (26%)	Arboretum, production forest, organic orchard

3.4.2. Soil texture

As former land use did, the difference in soil structure was one of the criteria on which the selection of food forest was based. There is a large variety of soil structures throughout the Netherlands. In the fall of 2019, the NMVB started analysing soil conditions in all food forests. Based on the amount of sand, silt and lutum particles in the soil samples, the food forests were divided in twelve categories, in a gradient from sand to clay (see *Table A1 & A2*). As for former land use, these soil textures were summarised in three classes: sand, loam and clay. The criteria of placing a certain soil in that specific category are shown in the third column of *Table 2*. These categorising is based on the NRCS soil texture triangle (Groenendyk *et al.,* 2015; see *Figure A1*). Again, the number of food forests in the second column is higher than the number of food forests included in this research (23 and 21 respectively), since one food forest had both clay-zones and loam-zones (food forest Roggebotstaete).

Table 2: Categorisation of 13 soil textures into three soil texture classes: Sand, Loam, and Clay. Classes has been made using the NRCS soil texture triangle.

Soil texture class	*Number of food forests*	*Criteria of this classes (see Figure A1 for underlying triangle)*	*Included soil textures*
Sand	9 (39%)	> 70% of soil are sand particles and < 15% of soil are clay particles	Sand, loamy sand
Loam	9 (39%)	< 52% of soil are sand particles and < 27% of soil are clay particles	Sandy loam, loam, silty loam, silt
Clay	5 (22%)	> 27% of soil are clay particles or > 20% of soil are clay particles and < 27% of soil are silt particles	Clay loam, sandy clay loam, silty clay loam, sandy clay, silty clay, clay

3.4.3. Food forest age

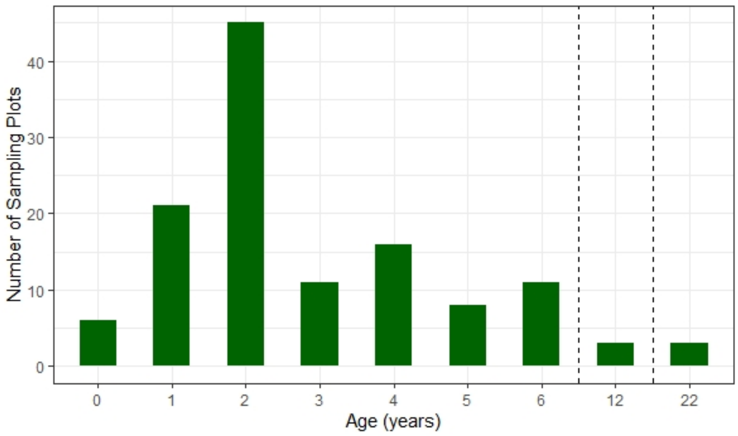


Figure 5: Distribution of food forest age across 133 sampling plots.

The third criteria on which food forests were selected is their age. By defining food forest age, one could analyse the carbon accumulation over time, as appointed in the main research question. There are several ways to define food forest age. In this research, food forest age is defined by the number of growing seasons since the first planting event. Since plants mainly grow in a specific growing season, which is roughly from April to August, it did not make a significant difference in growth whether the planting event took place in the fall or winter. Therefore, the number of growing seasons was preferred over food forest age in months. As shown in *Figure 5*, young food forests were over-represented compared to older food forests, though a broad age range was preferred by the NMPF. There are only two food forests older than six years in the Netherlands (Ketelbroek (12 years old) and Sualmana (22 years old)), which automatically caused the skewed shape.

Food forests could have also been divided in zone, based on their structural composition. In general, two major structural compositions were distinguished. In some of the food forests, trees and shrubs were planted in rows. These so-called *rational food forests* generally used these rows as it increased the usability of the food forest during harvesting. In other food forests, species were planted through each other and looked more like a natural forest. These food forests are commonly known as *romantic food forests*. Although the terms *rational* and *romantic* are applied in food forest community, the terms *alleys* and *dispersed* will used instead in this research, in order to increase comparability with agroforestry literature. So *romantic food forests* were called *dispersed food forests* and *rational food forests* were called *food forests with alleys*.

3.5. Data Analyses

After all measurements, calculation and categorisation of the variables was done, the data was analysed. First, the effect of the four explanatory variables (age, former land use, soil texture and structural composition) on carbon stock was visualised and examined using bivariate analyses. After this was done, the effect of multiple of these variables interacting with each other was inspected and carbon stock data was connected to soil data collected by the NMVB using a principal component analysis. The effect of management was checked by a study on the differences between zones within the same food forest. To conclude, the microclimate was analysed, and the development of succession was examined.

All analyses described below were performed using the statistical program R (version 4.0.2) and RStudio (R Core Team, 2020). Unless explicitly described, all analyses were performed using the carbon stocks per sample plot.

3.5.1. Bivariate analysis of explanatory variables

Both FLU and soil texture were classified in three classes and were analysed in the same way. First, group averages were visualised in boxplots. The package *ggplot2* was used in RStudio to derive these boxplots (Wickham *et al.*, 2016). Whether the differences in carbon stock between these groups was significant could be tested by a one-way analysis of variance (ANOVA) using the *car* package (Fox & Weisberg, 2019). To use an ANOVA legitimately, the data should be normal distributed, and the residuals of the data should be homogeneous. These assumptions were tested using a Shapiro-Wilk test for normality (Shapiro & Wilk, 1965) and a Levene's test for homogeneity (Levene, 1960) in R-packages *dplyr* and *car* respectively (Wickham *et al.,* 2020; Fox and Weisberg, 2019). If one or both tests rejected the null hypothesis of a normal distributed data and homogeneous residuals ($p<0.05$), the data was not directly suitable for a one-way ANOVA. After a logarithmic transformation of the data (*i.e.* of the response variable carbon stock), the assumptions were checked again. If the null hypotheses were accepted, the one-way ANOVA was performed with these log-data. If the null hypotheses were rejected again, the non-parametric Kruskal Wallis test was used as alternative for the one-way ANOVA (Kruskal & Wallis, 1952). These tests could only show whether there were significant differences between groups but did not conclude anything about which specific groups were significantly different. To find out, a Tukey's Honest Significance Test was performed (Abdi & Williams, 2010).

The effect of food forest age on carbon stock was analysed in a slightly different way. Since an increase of stored carbon over time was hypothesised, this relation was set out in a scatter plot. The regression of carbon stock over food forest age was also checked using an ANOVA. After the assumptions belonging to a one-way ANOVA were checked and log transformation was performed when required, the difference in carbon stock over time were tested. However, a default linear model (which is used in an ANOVA) might not be suitable. The accumulation of aboveground carbon in food forests over time is expected to be sigmoid (Birch, 1999; Dewar, 1990). Therefore, a sigmoid model was built-in in the linear model of carbon stock and food forest age. First, a polynomial regression model was made in RStudio using the *ggplot2* packages (Wickham, 2016), set with three folds (which provided a sigmoid curve). Using the nls() function in the *nlstools* packages (Baty, Ritz & Bath, 2015), a sigmoid model was made: *nls(Carbon ~ a/(1+exp(-b*(Age-c)))*. By setting parameters *a, b* and *c,* the appropriateness of the sigmoid model could be maximised. Parameter *a* described the saturation level of the sigmoid curve, parameter *c* is the age of food forest (x-axis) at half-saturation level and parameter *b* is influencing the slope of the model. By testing a correlation between the predicted and observed values, this appropriateness was checked (*corr.test*-function in R). The higher the significant correlation coefficient, the better the sigmoid model was following the relation between aboveground carbon stock and age of food forests.

3.5.2. Interaction between age, former land use and soil texture

After testing the relation between these variables and carbon stock of food forests individually, a model with interaction between these variables was made. For example, the increase of carbon stock over food forest age might be different for a food forest on clay than a food forest on a sandy soil. These interactions were tested using a generalised linear model. First, the assumptions of the default *glm* were tested (normality and homoscedasticity of the residuals). To do so, diagnostic plots and a density curve were made. In case of rejected assumptions, another distribution has been implemented in the *glm*. Again, the diagnostic plots were used to determine which distribution best fitted the data (see *Appendix 96*). Food forests with former land use category *forest* were excluded from the generalised linear model testing the interaction between age and soil texture, since FLU *forest* was expected to significantly affect this relation. Analyses of food forests without initial aboveground carbon stock only gave better insight in the effect of soil texture on the accumulation of carbon in food forests. Using the *summary*-function in RStudio (*car* package), the P-values of each variable were extracted from the generalised linear model to check whether slopes of the models were significantly different (*i.e.* to test whether the variables are interacting significantly).

To determine which *glm* was best predictive for the response variable (aboveground carbon stock), the *dredge*-function in the *MuMIn*-package was used (Barton & Barton, 2015). The generalised linear model with the lowest logarithmic likelihood was assumed to be best predictive. However, a simple model was preferred over a more complex model when the predictive capacity of the models was comparable. To correct the likelihood for the complexity of the model, the Akaike Information Criterion (AICc) was used (Akaike, 1974). Models with a delta AICc less than 2 were selected as best predictive. With this output, one could analyse which variables and interaction were most influencing the carbon stock of food forests. The numeric variable *Age* was averaged for the selected models (*MuMIn*-package), which could not be done for categorical variables. Since the difference in slope between different soil texture classes and former land use categories were interesting anyway, these *glm's* were used regardless of the AICc.

3.5.3. Relation between AGC stock and soil conditions

As mentioned before, the NMVB performed detailed soil analyses during the winter of 2019-2020. 18 out of the 21 food forests included in this research were analysed during that study. Therefore, it was interesting to connect the results of this research (i.e. the carbon stock of the food forests), with the most important results of that study. Out of the many variables tested, the following five were classified as most important explanatory variables of variance between food forests: amount of soil organic matter (SOM), acidity (pH), cation exchange capacity (CEC), total amount of available nitrogen (Ntot) and moisture content. Since the food forests were not divided into zones based before the soil samples were taken and analysed, the carbon stock will be analysed per food forest as well. For zoned food forests, the AGC stock was determined by averaging the carbon stocks of the zones. Food forests that were not analysed in the soil study were excluded from this analysis as well, making the total number of analysed carbon stocks set at 18.

To connect both studies, two analyses were performed. First, a principal component analysis (PCA) was performed, using the *FactoMineR* package (Le *et al.*, 2018). In this PCA, the correlation between all five soil variables and amount of aboveground carbon was visually analysed. As extra variable, the age of the food forest (in number of growing seasons since the first planting event) was added to this PCA. Former land use and soil texture could not be implemented in this PCA since these are categorical variables. The percentage of soil particles that were clay and sand were added to the PCA, to visualise the gradient of soil structure. Note that silt characteristics lay in between clay and sand and therefore they were not added to the PCA. The correlation between carbon stock and all those variables was tested using a correlation test. The assumptions for using the default Pearson correlation were tested by testing normal distribution of both variables and linearity between them. If the assumptions were not met, a rank-based Spearman correlation was used instead. The outcome of these correlation tests was visualised in a correlation matrix using the *corrplot*-function in the eponymous package (Wei & Simko, 2017). All non-significant correlations ($p>0.05$) were excluded from the correlation matrix.

3.5.4. Variation in carbon stock between food forests zones

As described before, some of the food forests were divided into zones based on their conditions. The conditions of these zones are assumed to be relatively similar, except for the one or two variables that distinguish these zones. Therefore, it was interesting to analyse the difference between those two zones. There were five reasons for the implementation of zones in the food forests selected by the NVMB: differences in management type, age, soil texture, height, and vegetation structure, as shown in *Appendix 1*. These differences were analysed using paired t-tests. In this way, the difference between two measuring points (*i.e.* zones) of the same food forests were analysed. Again, the assumptions of t-tests were checked (as described before), and when needed logarithmic transformation and/or non-parametrical alternative were used. These analyses gave insight in the effect of one specific variable on carbon stock, excluding many other variables, though the sampling size and therefore the reliability and power of the results was very limited.

3.5.5. Microclimate analyses

The measurements on microclimate included light availability (canopy closure), humidity and temperature. For the latter two, the analysis consisted of two parts. First, the difference between the temperature and humidity inside and outside the food forest was tested. Measuring points were the criteria on measuring the outside microclimate had not been met were excluded from this analysis. Since the data are measured in pairs (the outside temperature of food *forest X* had nothing to do with the inside temperature of food *forest Y*), a paired t-test was used. If the assumptions were not met, a Wilcoxon Signed Rank test was used as non-parametric alternative Wilcoxon, 1945). After the difference was examined, the relation between all three microclimate variables and carbon stock was tested. Therefore, the relative difference between the inside and the outside measurement was calculated for all zones (in percentages of the outside temperature). A linear model was composed to determine whether the regression of these variables over an increasing carbon stock was significant.

In addition to these analyses, a PCA with all three microclimate variables was made, using the *FactoMineR*-package (Lê, Josse & Husson, 2008). In contrast to the PCA with soil variables, this PCA was performed with all zones individually. Next to this, a correlation matrix has been made in the same way as the correlation matrix with soil variables.

3.5.6. Analysis of aboveground carbon stock measurements

At last, the relations between aboveground carbon stock and basal area, maximum height and canopy closure was analysed. The first two (basal area and maximum height) were both included in the allometric equations that were used to determine AGC stock. But the relation between AGC stock and both variables individually will be examined as well. In this way, one could investigate whether one of these variables is appropriate to estimate aboveground carbon stock of food forests. Canopy closure was not included in the allometric equations but is expected to be related to carbon stock. Therefore, the appropriateness of canopy closure as explanatory variable will be determined as well. A one-way ANOVA will be used to determine whether there is a regression and the strength of this regression for all three variables individually. The *dredge*-function (Barton & Barton, 2015) will be used in RStudio to determine which generalised linear model (containing one, two or all three variables) best declares variation in AGC stock.

Next to this, the relative amount of carbon stored in the shrubs was examined. Therefore, the amount of carbon stored in the shrubs has been set out to the amount of carbon stored in the trees. Moreover, the relative amount of carbon stored in shrubs (as percentage of total AGC stock) will be plotted against the total AGC stock. Linear models will be used to determine significance and strength of both regressions. These analyses could answer the question whether the amount of carbon stored in shrubs changes over time, or whether this is a fixed percentage, in order to improve the accuracy and/or accessibility of the methods of the NMPF next year.

4. Results

In this section, the most important results of this study are displayed and described. The results of assumption tests and supplementary results are depicted in the appendix.

4.1. Aboveground carbon stock over food forest age

To get insight in the influences on the four explanatory variables measured in this research, the relation between aboveground carbon stock and these variables were analysed. First, bivariate statistics were applied to these explanatory variable each seperately. After that, any potential interaction between those variables were analysed (multivariate analyses), for example to test whether there was a difference in carbon accumulation over time between soil texture classes.

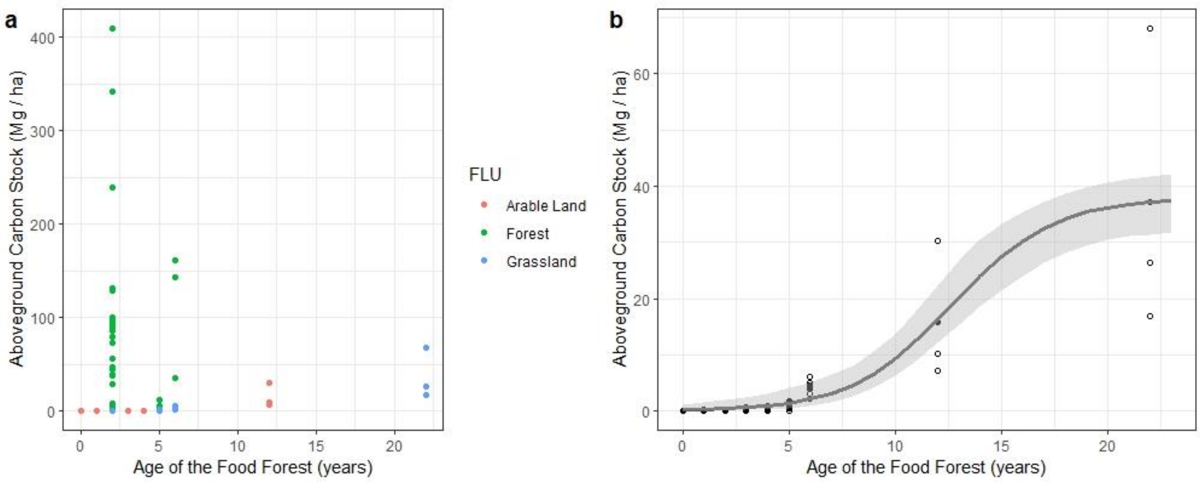


Figure 6: Food forest aboveground cabon stock versus food forest age. Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}, age is expressed in years. **a)** Aboveground carbon stock versus food forest age for all food forests. Filled dots representate sampling plots, coloured by their former land use category. n = 130. **b)** Sigmoid model curve, including food forests with FLU *arable land* and *grassland*. Open dots respresent all sampling plots, n = 95; filled dots represente means per zone, n = 29. Sigmoid model is significantly correlated with observed mean values (coëfficient = 0.99, $p < 2.2e\text{-}16$). Parameters of the sigmoid model determined at a = 40, b = 1, c = 5. *Exported from RStudio.*

Carbon stock was not increasing over time, based on all food forests (see *Figure 6*). However, since some of the food forests already had a notable aboveground carbon stocks in trees (*i.e.* food forests with former land use *forest*), these data did not provide a clear picture of carbon accumulation. The effect of food forest age on carbon stock therefore was determined for food forests without any initial carbon stock in trees only (*i.e.* food forests with former land use categories *grassland* and *arable land*; see *Figure A6*). The relation between aboveground carbon stock and food forest age, seemed to be a sigmoid curve. The best fitting sigmoid curve, obtained using the *nls*-function in R (see *Methods*) was:

$$Sigmoid\ Model = nls(Aboveground\ Carbon\ Stock \sim \frac{40}{e^{-(Age-5)}} \qquad \textit{(Equation 8)}$$

In this model parameters a, b and c were determined at 40, 1 and 5 respectively. A was determined at 40, since aboveground carbon stock was saturating at 40 Mg ha^{-1}; b and c were chosen since these values maximised the fitting of the sigmoid model. The coefficient of correlation between predicted and observed values was 0.99, and the correlation was significant ($p < 2.2e\text{-}16$). The relation between aboveground carbon stock and food forest age is therefore proven sigmoid. When a sigmoid model was made based on the individual sampling plots, the correlation was still significant, albeit with a smaller correlation coefficient (0.84; $p < 2.2e-16$). See *Figure A6* for additional graphics, displaying the sigmoid curve.

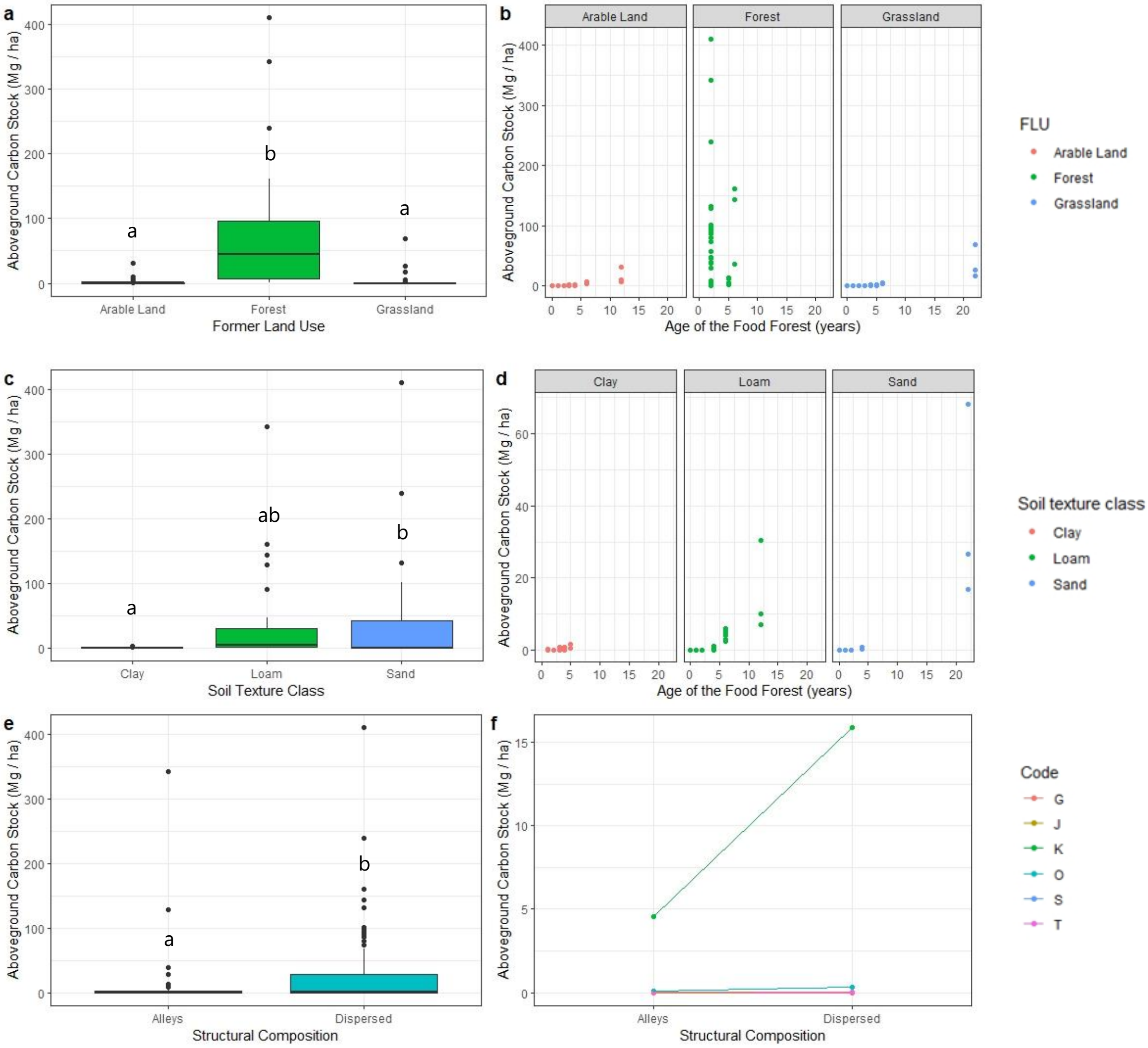


Figure 7: Food forest aboveground carbon stock versus categorical variables soil texture, former land use and structural composition. Scatter points representate sampling plots. Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}, age is expressed in years. **a)** Boxplot of the aboveground carbon stock per former land use category. n = 130. Carbon stock in FLU *forest* was 62.8 and 64.8 Mg ha^{-1} higher than *arable land* and *grassland* ($p < 1e-05$). **b)** Regression of aboveground carbon stock over age for three categories of former land use for all food forests. n = 95. **c)** Boxplot of the aboveground carbon stock per soil texture class. n = 130. Carbon stock on sandy soils was 34.01 Mg ha^{-1} higher than on clay soils ($p = 0.01$). **d)** Regression of abovegroud carbon stock over age for three soil texture classes for all food forests with FLU *grassland* and *arable land*. n = 95. **e)** Boxplot of the aboveground carbon stock per structural composition. n = 130. Dispersed food forests had a higher aboveground carbon stock than food forests with alleys (difference = 2.2 Mg ha^{-1}; $p = 0.03$). **f)** Lineplot of food forests with both structural compositions. Strata means are shown. n = 12. Paired t-test showed significant higher carbon stock in dispersed food forests ($p = 0.02$). *Exported from RStudio.*

4.2. Aboveground carbon stock per former land use, soil texture and structural composition

The other three explanatory variables included in this research were all categorical. Therefore, the analysis of the influence of these variables, were graphically presented using boxplots of the mean carbon stock per category (see *Figure 7*). For all three analyses, the assumptions of the tests were not met and non-parametric statistical tests (Kruskal-Wallis significance test and Mann-Whitney U test) were used (see *Table A4* for the outcomes of the assumption tests). Carbon stock was significantly higher in food forests with former land use *forest*, than food forests with FLU-categories *arable land* and *grassland*, with a difference of 64.8 Mg ha^{-1} and 62.8 Mg ha^{-1} respectively ($p < 1e\text{-}05$). No difference was found between FLU-category *arable land* and FLU-category *grassland* ($p = 0.975$). Aboveground carbon stock of food forests was also significantly different between soil texture classes ($p < 0.001$). Carbon stock on sandy soils was estimated at 34.01 Mg ha^{-1} higher than clay soils ($p = 0.01$), but no difference was found between sand and loam or clay and loam ($p = 0.31$ and $p = 0.27$ respectively). A significant difference was also found between the two ways of structuring the food forest. Dispersed composed food forests had a higher carbon stock than food forests with *alleys*, both in an analysis of all food forests ($p = 0.03$) and in an analysis of paired measurements (*i.e.* food forests that contained both structural compositions; $p = 0.02$).

The first multivariate analyses have been performed on the interaction between food forest age and soil texture, and the interaction between food forest age and former land use (see *Figure 7b,d*). A generalised linear model with *Gamma*-distribution has been used in both examinations, since this maximised the likelihood of the *glm* (note that the assumptions of the default glm (*Gaussian* distribution) were not met, see *Figure A2-A5*). Food forests with FLU *grassland* showed a significant smaller slope than food forests with FLU *arable land* (0.65 ± 0.07 and 0.44 ± 0.08 respectively; $p = 0.01$). The slope of carbon stock over age for food forests with FLU *forest* was also significantly smaller than the slope of former *arable lands* ($p < 0.001$). The direction of this slope had a negative estimate, indicating a decreasing carbon stock over time in these food forests (-0.10 ± 0.17) The effect of soil texture on this accumulation was examined in the same way. Food forests on clay had a significant higher slope than food forests on sand, with estimated slopes of 0.74 ± 0.19 and 0.33 ± 0.19 respectively ($p = 0.03$). Carbon stock accumulation was not different from both sand and clay soils in this research (both $p > 0.05$), although the slope was expected to be higher than the slope on sand soils (slope of 0.67 ± 0.21).

Table 3: The output of the *dredge* analyses for selecting the best explanatory generalised linear models. Degrees of freedom (df) indicates model complexity; less degrees of freedom indicates a simpler model. Log likelihood (LogLik) indicates the explanatory capacity of the model; a lower LogLik indicates a better explaining model. Akaike Information Criterion (AICc) ranks models based on both explanatory capacity and complexity; the model with the lowest AICc is the best model. Delta AIC shows the difference in AICc between a model and the best scoring model; all models with a delta < 2 were selected.

Model	df	LogLik	AICc	Delta AIC	Adj. R^2	p
glm(Carbon ~ FLU + Comp + FLU*Comp)	7	-653.99	1322.94	0.00	0.35	1.49e-11
glm(Carbon ~ FLU + Comp + Age + FLU*Comp)	8	-653.37	1323.99	1.05	0.35	3.49e-11
glm(Carbon ~ FLU + Comp + Age + FLU*Comp + Age*Comp)	9	-652.67	1324.91	1.97	0.35	6.94e-11

Although these models described differences in carbon stock accumulation between soil texture and FLU, many more *glm's* describing variety in carbon stock could be made. With the *dredge*-function in R, all possible *glm's* have been made, including all four explanatory variables age, soil texture class, FLU class and structural composition and interactions between these variables. Three models had a delta AIC of less than 2, and therefore assumed to have the strongest explanatory power. As shown in *Table 3*, all these models contained variables FLU, structural composition, and an interaction between these two. The third model, additionally including variable age and an interaction between age and structural composition, had the best predictive capacity (*i.e.* the lowest absolute log likelihood) of these three options. However, this model is more complex (*i.e.* more degrees of freedom), which moderated the AIC. The explanatory variable soil texture was not included in one of these best predictive *glm's*.

4.3. Food forest microclimate

As described before, food forests were expected to have an influence on the microclimate of their environment. Both temperature and humidity were measured inside and outside the food forest, and the means of the measurements are displayed in *Figure 8*. For both analyses, all assumptions were met (see *Table A4*) and a t-test for two paired samples was used (since the outside and inside measurements belong to the same food forest). The inside air temperature was significantly lower than the outside temperature, with a mean distance of 10.14 ± 2,67 °C ($p < 0.001$). The humidity of the air was higher inside the forest than outside of it, with a mean difference of 12.03 ± 4.68 % saturation ($p < 0.001$). De development of these differences as carbon stock increases is also shown in *Figure 8*. The relative differences of both temperature and humidity were not significantly changing over carbon stock, based on a regression analyses ($p = 0.71$ and $p = 0.58$ respectively).

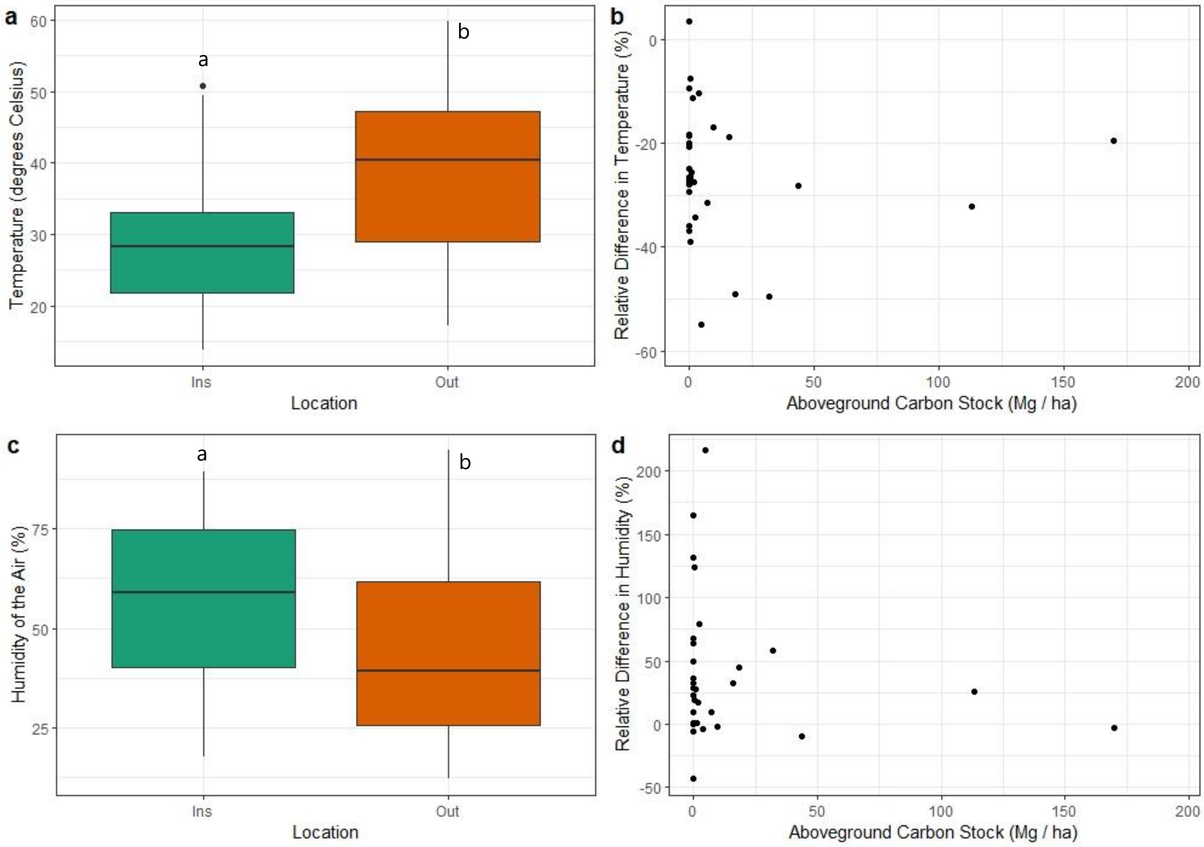


Figure 8: The relation between food forest vegetation and temperature and humidity. Scatter points representate zones, n = 31. No data were obtained in foodforests B, C, F, M and P. Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}, humidity in percentage saturation, temperature in degrees Celsius and relative differences in percentages between inside and outside measurements. T-test for two paired samples were used to determine differences between inside and outside measurements, one-way ANOVA was used to determine regression of temperature over aboveground carbon stock, the non-parametric alternative Kruskal-Wallis significance test was used for humidity regression. **a)** Boxplot of the mean temperature inside and outside the food forest. Mean outside temperature was 10.14 ± 2,67 °C higher than mean inside temperature ($p < 0.001$). **b)** Percentage difference between inside and outside temperatures for all zones versus the carbon stock of this zone. No significant regression was found ($p = 0.71$). **c)** Boxplot of the mean humidity inside and outside the food forest. Mean inside humidity was 12.03 ± 4.68 % higher than mean outside humidity ($p < 0.001$). **d)** Percentage difference between inside and outside temperatures for all zones versus the carbon stock of this zone. No significant regression was found ($p = 0.58$). *Exported from RStudio.*

The relation between carbon stock and microclimate of a food forest is also analysed using Spearman correlations (see *Figure 9*). Similar to the outcome of the regression analysis, no significant correlation was found between aboveground carbon stock and temperature of humidity ($p > 0.05$ for both correlations). In a correlation matrix without soil variables, where number of measurements increased from 12 to 31 (since these data were available per zone), these correlations were still not significant ($p > 0.05$, see *Figure A7*). Although it was not significant, the correlation coefficient of humidity versus carbon was higher than the correlation coefficient of temperature versus carbon ($|0.47|$ to $|0.03|$). These findings corresponded with the output of the principal component analysis, where humidity was on the same axis with carbon stock (*i.e.* in opposite direction), while temperature and carbon stock were not on the same axis (*i.e.* more or less perpendicular to each other).

4.4. Aboveground carbon stock versus soil conditions

In this paragraph, the connection between aboveground carbon stock and soil conditions has been made. First, a principal component analysis (PCA) of the data has been performed (*Figure 9a*). Next to the soil variables (total amount of nitrogen, cation exchange capacity, soil organic matter, acidity, moisture) and aboveground carbon stock, all other numeric variables are added to the PCA (maximum height of the tree layer, age, difference in temperature and difference in humidity). All variables were implemented at food forest level since the soil data was not available per zone or plot. All variety within a food forest was not analysed in this way. According to the PCA, one could argue that all five soil indicators are highly correlated with each other, and with the amount of lutum (clay) particles in the soil. The sand arrow was located in the opposite direction of these six arrows, indicating a highly negative correlation between the amount of sand particles and these six variables. Soil organic matter (SOM) seemed to be the soil variable most correlated with carbon stock since the SOM-arrow is the closest to the Carbon-arrow. No significant regression between AGC stock and SOM content was found ($p = 0.20$, see *Figure A9*). One could argue that there is an abiotic axis from sand to moisture, and a biotic axis from carbon to humidity. Clusters of food forests became visible, having food forests on a sandy, nutrient poor soil (E, H, I, J, L); food forests on a clay, nutrient rich soil (D, G, O Q); and food forests with a relatively large carbon stock (A, P; although they are quite different in soil conditions). See *Appendix 1* for information about these codes and the corresponding food forests.

These assumed correlations can be checked using a correlogram. Since the assumptions for testing correlations were not always met, a rank-based Spearman correlation function was used. A significant correlation between the five soil variables was confirmed, just as the significant negative correlation with the amount of sand particles (all $p < 0.05$; see *Table A5* for all specific p-values). Carbon stock was not significantly correlated with one of the soil variables (all $p > 0.05$). From all five soil variables, carbon stock was most correlated with SOM (adj. $R^2 = 0.52$) and less correlated with CEC (adj. $R^2 = 0.09$), although these correlations were not significant as mentioned. In *Figure A8,* a heatmap is visualised, where clusters of most correlating variables are displayed. The two distinguished clusters are: 'pH, Lutum, SOM, CEC, Ntot and Moisture' and 'Carbon, Maximum Height, Canopy and Age'.

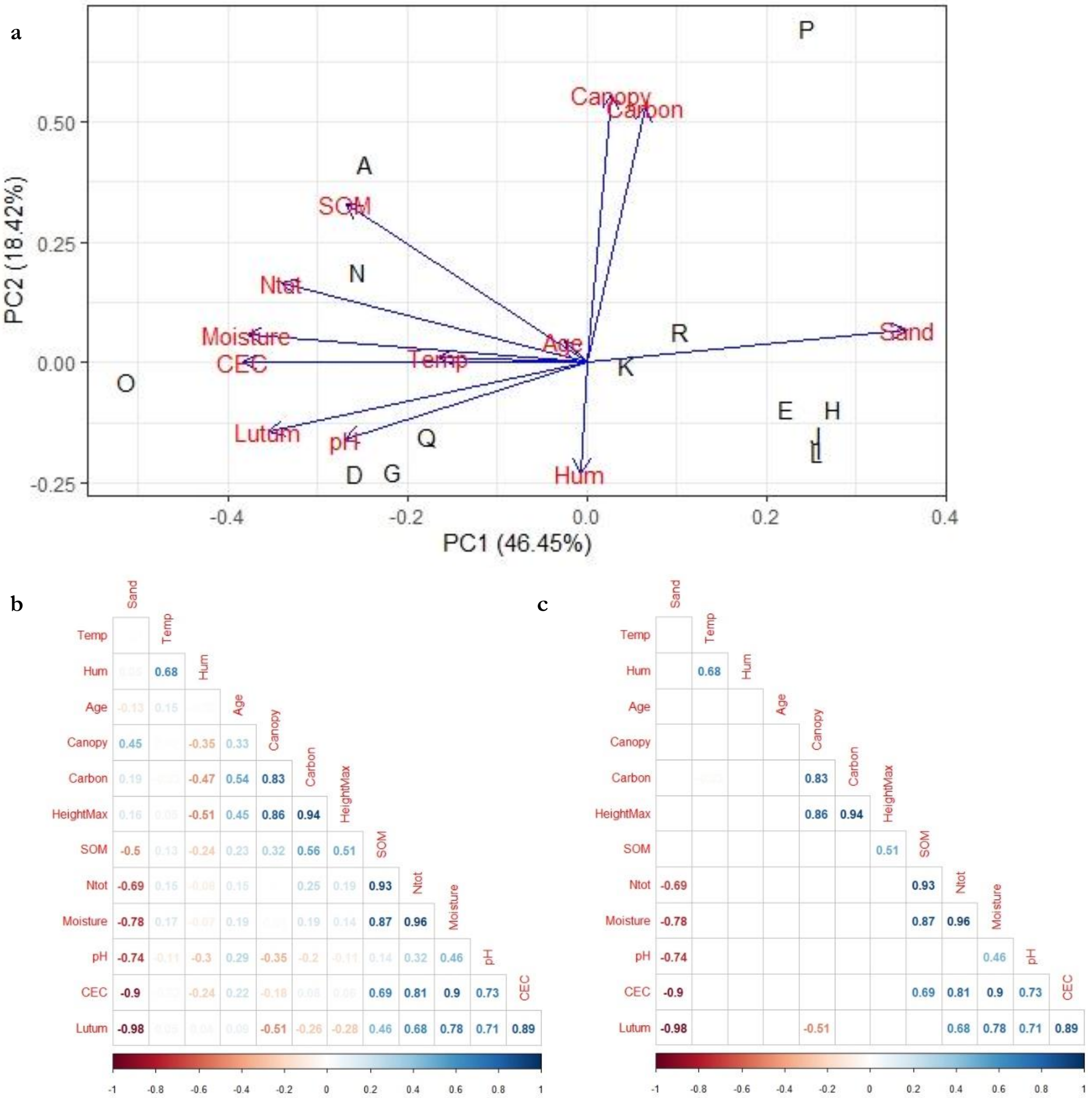


Figure 9: Relation between response variable carbon stock and explanatory variables age, relative difference in temperature (Temp), relative difference in humidity (Hum), percantage canopy closure (Canopy), soil organic matter (SOM), cation exchange capacity of the soil (CEC), moisture availabilty of the soil (Moisture), total amount of available nitrogen (Ntot), acidity (pH), percentage of clay paticles (Lutum) and percentage of sand particles (Sand). Input information of all three analyses was the mean per food forest, n = 12. **a)** Principal component analysis (PCA), codes of the food forests as scatter points. **b)** Correlogram showing adjusted R^2 values of all possible correlations, based on the Spearman, rank based correlation method. Positive correlations in blue, negative correlations in red. **c)** Correlogram showing adjusted R^2 values of all significant correlations ($p < 0.05$) *Exported from RStudio.*

4.5. Elements of aboveground carbon stock calculations

In order to determine whether it is possible to estimate aboveground carbon stock based on easier, less complex measurements, the regression between AGC stock and three variables are analysed: closure of the canopy, maximum height and basal area. Next to this, the importance of measuring shrub carbon stock is examined. As shown in *Figure 10,* a significant linear regression was found between canopy closure and AGC stock ($p < 0.05$). However, this regression model only explained 33% of the variation in canopy closure (adjusted $R^2 = 0.33$) and the relation seemed exponential instead of linear (see *Figure 10a*). The turning point is located at an AGC stock of ~ 15 Mg ha^{-1} (see *Figure A13*). After a logarithmic transformation of the carbon stock (see *Figure 10b*),

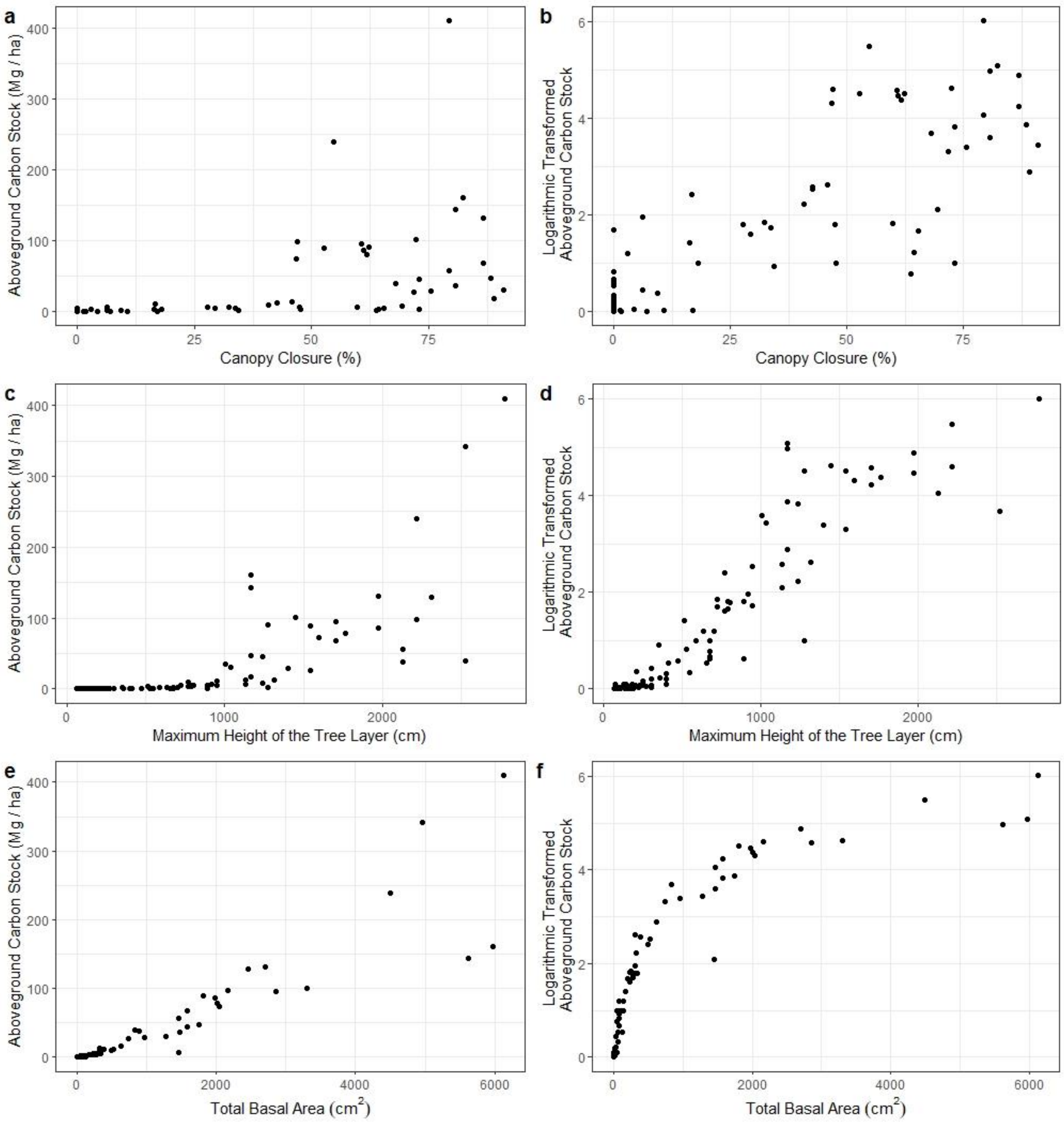


Figure 10: Canopy closure, maximum height of tree layer and basal area versus aboveground carbon stock. Scatter points representate sampling plots. Canopy closure data was obtained using Canopy Capture application (Patel, 2018); no canopy closure data was obtained in foodforest C (MijnStadstuin, Amsterdam). Trees and shrubs were be considered to be one layer. Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}, canopy closure in %, maximum height in centimeters. One way ANOVA's were used to test regressions. **a)** Aboveground carbon stock versus closure of the canopy; significant regression was found with an adjusted R^2 of 0.33 ($p < 0.001$). n = 121. **b)** Logarithmic transformed aboveground carbon stock versus closure of the canopy; significant regression was found with an adjusted R^2 of 0.76 ($p < 0.001$). n = 121. **c)** Aboveground carbon stock versus the maximum height of the tree layer; significant regression was found with an adjusted R^2 of 0.56 ($p < 0.001$). n = 130. **d)** Logarithmic transformed aboveground carbon stock versus the maximum height of the tree layer; significant regression was found with an adjusted R^2 of 0.87 ($p < 0.001$). n = 130. **e)** Aboveground carbon stock versus closure of the canopy; significant regression was found with an adjusted R^2 of 0.82 ($p < 0.001$). n = 130. **f)** Logarithmic transforemd aboveground carbon stock versus closure of the canopy; significant regression was found with an adjusted R^2 of 0.69 ($p < 0.001$). n = 130. *Exported from RStudio.*

the relation between both variables was linear. In this model, there is a significant regression between canopy closure and carbon stock, with an adjusted R^2 of 0.76. This log-transformed model better described the relation than the non-transformed model, confirming the logistic regression between carbon stock and canopy closure. As shown in *Figure 9,* canopy closure was also significantly correlating with AGC stock, with a correlation coefficient of 0.86 ($p < 0.05$).

The relation between carbon stock and the maximum height of the tree layer is comparable. Again, there is a relation between the two variables: carbon stock increases if maximum tree height increases. A significant linear regression was found, with an adjusted R^2 of 0.56 ($p < 0.001$). After logarithmic transformation of the carbon stock, a more linear looking relation was found (see *Figure 10d*). The linear regression of this model declared over 87% of the variation between maximum height of the tree layer and carbon stock, confirming that this model better explained the relation than the non-transformed model. The relation between maximum height of the tree layer and carbon stock was therefore significant and exponential. The correlation matrix has also shown a significant correlation between the maximum height of tree layer and AGC stock, with a Spearman's correlation coefficient of 0.96 ($p < 0.05$); making maximum height the variable most strictly correlating with aboveground carbon stock (*i.e.* variable with the highest correlation coefficient, see *Figure 9*).

Next to tree height, the allometric equations required stem diameter measurements. Stem diameter is converted to total basal area per sampling plot. In contradiction to both canopy closure and maximum height of the trees, basal area showed a linear regression with AGC stock (no logarithmic transformation was needed). This regression declared up to 82% of variation in carbon stock ($p < 0.05$; see *Figure 10e*). As shown in *Figure 10f*, a logarithmic transformation did not improve linearity nor the explanatory capacity of the regression (Adj. $R^2 = 0.69$). Basal area was declaring less variation in AGC stock than maximum height of the tree layer did (82% and 87% respectively), although a logarithmic transformation was required for the latter.

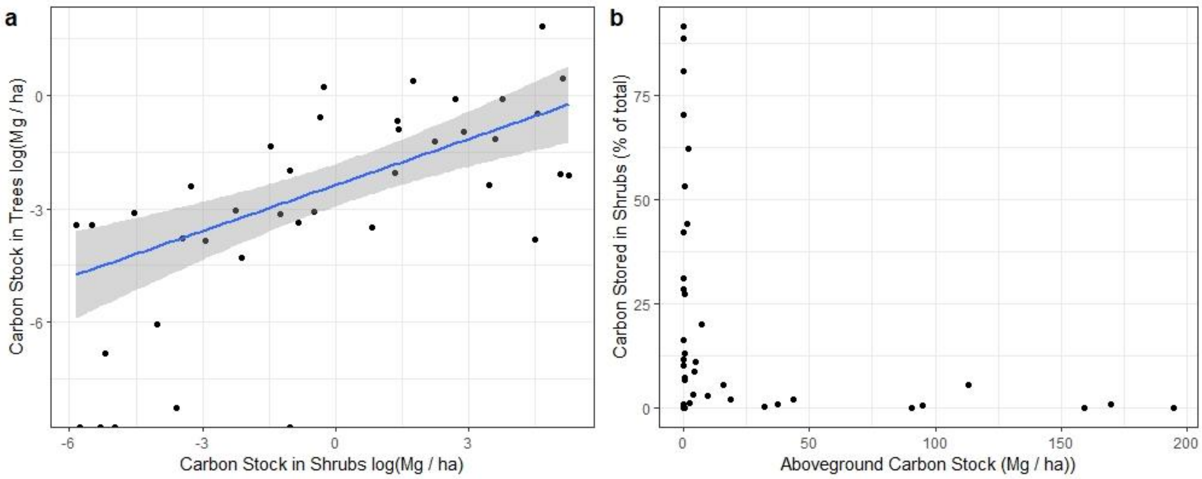


Figure 11: Amount of carbon stored in shrubs. Scatter points representate sampling plots (n = 130). Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}. Spearmon, rank-based correlation methods were used to test correlations, one way ANOVA's were used to test regressions. **a)** The amount of carbon stored in trees versus the amount of carbon stored in shrubs, both logaritmic transformed; significant correlation of 0.55 was found ($p < 0.001$). **b)** Relative amount of carbon stored in shrubs (as percentage of total aboveground carbon stock) versus aboveground carbon stock aboveground carbon stock; significant regression was found with an adjusted $R^2 = 0.09$ ($p < 0.04$). *Exported from RStudio.*

To determine whether it is accurate to determine aboveground carbon stock by measuring only one variable, an analysis of the possible generalised linear models was performed. Using the *dretch*-function in RStudio, the best predictive models were selected out of all models including explanatory variables canopy closure, maximum height of the tree layer and basal area (see *Table A8*). The *glm* with all three variables was determined to be the best explanatory model (lowest AICc) and explained 99% of variation in aboveground carbon stock. The second-best predictive model was the model without canopy closure as explanatory variable, indicating that out of these three variables, canopy closure is the least accurate one. The models including either maximum height or basal area had a large delta AIC and were therefore describing AGC stock significantly worse than the model including both variables. A *glm* with maximum height as only explanatory variable described 55% of variation in AGC stock, while the *glm* with basal area as only variable explained 83% of AGC stock.

In order to determine whether shrub carbon stock measurements were necessary in further research, the relative amount of carbon stored in shrubs was analysed. First, the carbon stock in shrubs has been plotted against carbon stock in trees (see *Figure 11*). After logarithmic transformations, a linear relation was found (see *Figure A14* for non-transformed graphs). A significant correlation with a coefficient of 0.55 was found for this relation ($p < 0.001$). This indicated that carbon stock in shrubs is larger when carbon stock in trees is larger. Next to this, the amount of carbon stored in shrubs as percentage of total carbon stock has been analysed (see *Figure 11b*). In young food forests, the shrub carbon stock is relatively high, and it seemed to decrease over total aboveground carbon stock. A significant negative regression was found, with a slope of -0.64 ± 0.29 ($p < 0.04$). Note that the percentage of variance in data explained by this regression is low (adjusted $R^2 = 0.09$).

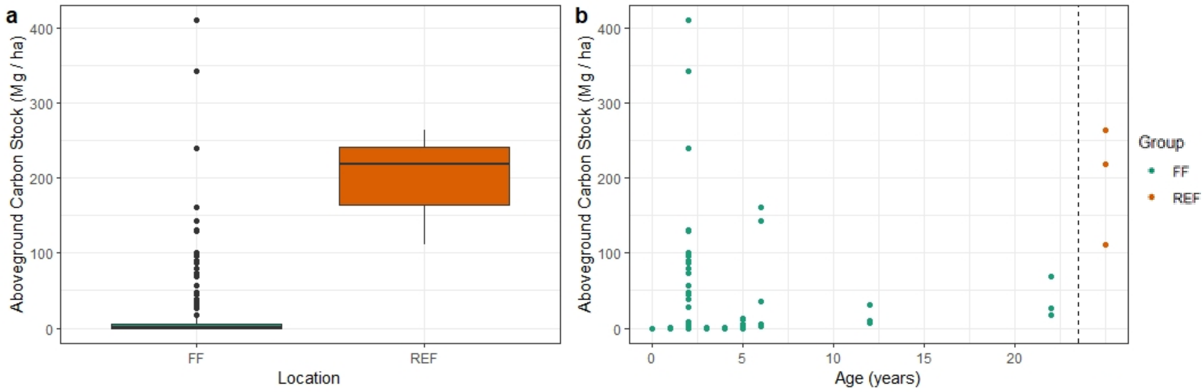


Figure 12: Aboveground carbon stock of food forests compared to aboveground carbon stock of more natural food forests. FF = food forest; REF = reference (natural forest). Boxplots were made using carbon stocks per sampling plots. Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}. **a)** Boxplots of the mean abovevround carbon stock for food forests (n = 130) and reference forests (n = 3). ACS was higher in reference forests, with 95% confidence interval = [110.3;263.4]; difference was significant ($p = 0.005$) based on Mann Whitney U Test **b)** Scatter plot of aboveground carbon stock versus age of the (food) forest. Red scatters representate sampling plots in food forests (n = 130), blue scatters representate sampling points in reference forests (n = 3).

4.6. Food forests versus natural forests

In principle, food forests are designed to function like a natural forest, including all ecosystem functions natural forests have. According to this hypothesis, carbon stocks of food forests should be comparable to carbon stocks of natural forests in similar circumstances. Three natural forests, one per soil texture class, were selected and aboveground carbon stock was measured, using one sampling plot each. As displayed in *Figure 12* mean aboveground carbon stock in reference forests was significantly higher than mean AGC stock in food forests. This difference was estimated at 216.59 ($p = 0.005$). Food forests with FLU *forest* were also significantly different from reference forests, although the difference was smaller (Estmiated at 135.7; $p = 0.01$; see *Figure A11*). At this moment, AGC stock in food forest is not at the same level as AGC stock in natural forets in the Netherlands.

Table 4: Aboveground carbon stocks found in this research compared to other food forests, agroforestry systems and natural forests. Values marked with * are not an average, but based on one study site. Number of measurements is only given for data obtained in this study.

Category	Location	Number of measurements	Mean age (years)	Mean AGC stock (Mg CO_2 ha^{-1})	Source
Food forests measured in this research	Food forests	130	3.43	21.80	-
	Food forests with FLU *forest*	35	2.84	66.16	-
	Food forests with FLU *arable land* or *grassland*	95	3.58	2.31	-
	Food forest Sualmana, the Netherlands	3	22*	37.16*	-
Other food forests and agroforestry systems	Agroforestry Research Trust food forest in Devon, United Kingdom	-	23*	34.53*	Schafer, Lysák & Henriksen, 2019
	Tropical home gardens and agroforestry	-	21	61.5	Rooduijn *et al.* (unpublished data)
Natural forests	Natural forest in tropics (secondary)	-	20	60.9	Rooduijn *et al.* (unpublished data)
	Natural forest in temperate climates	-	20	45	Rooduijn *et al.* (unpublished data)
	Natural forest in the Netherlands	-	50	59	Nabuurs & Mohren, 1993
	Natural forest in the Netherlands		max. 120	up to 200	Sikkema & Nabuurs, 1994
	Reference forests	3	unknown	197.5	-

AGC stock is also compared to data obtained from other scientific studies. As shown in *Table 4*, the 21 food forests included in this study had a mean aboveground carbon stock of 21.8 Mg ha^{-1} and a mean age of 3.4 years. Since the mean age was that young, it was hard to compare all these forest to natural forests, other temperate food forests and tropical equivalents. The oldest food forest, Sualmana, had an age of 22 years, which was only slightly different from the food forest in Devon, UK. As far as known, these are the oldest two temperate food forests in the world. The aboveground carbon stock of Sualmana was comparable to that of the food forest in Devon, 37.2 and 34.5 Mg ha^{-1} respectively (Lehmann *et al.,* 2019). These values were notably lower than the mean AGC stock in tropical home gardens with a more or less similar age (61.5 Mg ha^{-1}), assuming that carbon stock is built up quicker in tropic than in temperate food forests or agroforestry systems. A same ratio was found between mean AGC stocks in tropic and temperate natural forests, which were determined at 60.9 and 45 Mg ha^{-1} respectively at an age of 20 years (Rooduijn *et al.*, unpublished). Sualmana food forest had a lower aboveground carbon stock than a temperate natural forest with a similar age, albeit a small difference (37.2 to 45 Mg ha^{-1}). The reference forests included in this research had a mean aboveground carbon stock of 197.5 Mg ha^{-1}, which is significantly higher than a natural forest in the Netherlands with an age of 50 years (determined at

59 Mg ha^{-1} by Nabuurs & Mohren (1993)). However, aboveground carbon stock of natural forests can accumulate up to 200 Mg ha^{-1} after 120 years (Sikkema & Nabuurs, 1994). Although the age of the reference forests was unknown, it was unlikely that they were more than years old. All mentioned comparisons were not tested statistically, since every study had different measurements and methodologies. It could therefore only give some insight, rather than provide firm conclusions.

5. Discussion

The most important results were presented and described in the former paragraph. Food forests planted in former forests had a higher AGC stock than the other food forests, and food forests on sandy soils had a higher AGC stock than food forests on clay soils. Note that the mean age of food forests on sand was higher than food forests on clay as well. If FLU-category *forest* was left out, a sigmoid relation between aboveground carbon stock and the age of food forests has been found. The structural composition of food forests did also describe variation, as a significant higher AGC stock was found in dispersed food forest compared to ones with alleys. The accumulation rate of carbon was higher in food forests on former arable lands than those on former grasslands, and higher in food forests on clay than those on sand. Focussing on the microclimate of food forests, significant lower temperatures and higher humidity were found inside food forests compared to the outside conditions. These differences were not increasing when food forests have a bigger carbon stock. We have also seen that food forests were significantly correlating with canopy closure, the maximum height of their trees and basal area, from which maximum height was explaining most variation in AGC stock. From all included soil variables, the amount of soil organic matter was most correlated to AGC stock, although this correlation was not significant. At last, food forests included in this research did not yet reach the AGC stock values found in natural forests, but the oldest one (Sualmana) had a comparable AGC stock as a food forest examined in another study with similar age. The reference forests measured in this study have shown relatively high carbon stocks compared to mean AGC stocks of natural forests in literature.

All mentioned results will be discussed in this paragraph, complemented with the most important limitations of this research and recommendations for further research.

5.1. Aboveground carbon stock over food forest age

Food forest aboveground carbon stock was increasing as food forests were getting older. As hypothesised, this relation could be predicted by a sigmoid model, but one could discuss whether a saturation of AGC stock at a 40 Mg ha^{-1} is likely (Dewar, 1990; Birch 1999). This saturation level was reached at an age of less than 25 years. At this age, forests are still growing, and carbon stock is still developing (Sikkema & Nabuurs, 1994; Lee, McCarl & Gillig, 2005; Kauppi *et al.,* 2010. The biggest constraint of the sigmoid model is the lack of relatively old food forests and therefore the great dependence on one food forest (food forest Sualmana, located in Swalmen) concerning the curve of the regression. Food forest Sualmana is likely to have a smaller carbon stock than expected at an age of 22 years old, because of the nutrient poorness of its soil and the impoverishing management that has been executed. Although other studies have confirmed the negative effect of soil nutrient poorness on carbon storage in natural forests (Basset, 1994; Sullivan *et al.,* 2015), these studies have not been performed in food forests yet. The only relatively old temperate food forest (in Devon, UK) had a comparable AGC stock (Schafer, Lysák & Henriksen, 2019) as Sualmana. This food forest has a relatively nutrient poor soil as well, which could declare its unexpectedly small carbon stock (Lehmann *et al.,* 2019; Schafer, Lysák & Henriksen, 2019). Although mean carbon stock in temperate food forests of 22 years is expected to be higher than in Sualmana and Devon, this could not be confirmed due to the lack of other food forests with this age. In conclusion, the sigmoid curve could be realistic, but the great uncertainty of carbon stock in older temperate food forest should be borne in mind.

The accumulation rate of AGC was different in three categories of former land use *forest, arable land,* and *grassland.* A negative relation between carbon stock and age was found for food forests with FLU *forest.* This could be explained by the large variety of initial conditions between these food forests. For example, it is different to compare a former production forest and a former orchard. The former orchard will show a smaller carbon stock than the former production forest, even when it is several years older (Lugato *et al.,* 2014; Cao, Valsta & Mäkelä, 2010), which could result in a negative curve of AGC stock over time. The development of carbon stock should be monitored within one food forest to prevent the results for being affected by this variety in initial conditions. But the negative relation between age and AGC stock could also be caused by the fact that some of the food forests were thinned out, in order to enhance light availability for understory and clear the way for planting crop species. AGC stock will decrease in the first years but is expected to increase thereafter when planted trees and shrubs started growing. In fact, carbon stock accumulation rates are expected to be even higher in FLU *forest* than on former arable lands and grasslands. In general, soils of forests have a higher nutrient richness and water storage capacity than soils of grasslands and arable lands (Johnson & Wedin,1997; Billings, 2006; Evrendilek, Celik & Kilic, 2004). This could lead to a higher plant growth rate and therefore higher carbon accumulation rates (Bassett, 1964; Sullivan *et al.,* 2015). Moreover, the environmental circumstances of a forest provide more sun and wind protection than the circumstances on an arable land or grassland. On the other side, competition for water and nutrients will be heavier in forests than in arable lands and grasslands (Nambiar & Sands, 1993; Coomes & Grubb, 2000). Whether carbon accumulation rate is higher in former forests than in former arable lands and grasslands could therefore not be determined yet. At least a negative slope of AGC accumulation in former forests is implausible in the longer term.

5.2. Effect of categorical variables FLU, soil texture and structural composition

As hypothesised, AGC stock was significantly different between FLU-category *forest* and the other two FLU categories. Because food forests on former forests had a large initial aboveground carbon stock, this result is not astonishing. A larger carbon stock on former grasslands compared to former arable lands was not found, although this was hypothesised due to its microbial community and the higher nutrient and water availability this provides (Schulte *et al.,* 2005; Girvan *et al.,* 2004; McLauchlan, Hobbie & Post, 2006). The absence of this difference could not be explained by the mean age of both categories since the mean age of former grasslands was not higher than that of former arable lands. It might be possible that the differences in initial soil conditions between former grasslands and arable lands do not influence AGC accumulation at all. This corresponds to the absence of significant differences in soil conditions between both FLU categories in the study of De Groot (2020). Note that the effect of former land use could remain visible in the soil for decades (Bissett, 2011; Callaham, 2006). In this study, food forests are allocated to a FLU category based on their latest land use. But it is way more complex, and it might matter whether a food forest was an arable land for 5, 10, or 50 years (Bisett, 2011; Callaham, 2006). Added to this, not every arable land is the same. In short, a more detailed examination of former land uses should answer the question whether there is no difference in AGC accumulation between former arable lands and grasslands, or whether this absence is caused by the simplified analysis. An expansion of the dataset could help to answer this question, since the long-term effect of FLU types could better be analysed.

As expected, a significant difference in aboveground carbon stock was also found between soil texture classes sand and clay. Due to the higher availability of nutrients and moisture in clay soils, a higher AGC stock was expected in clay soils than in sandy soils (Bassett 1964; Brown, 2007; Sullivan *et al.,* 2015). However, the observed difference was contrary to this and a higher AGC stock on sand than on clay was found. Although carbon stock was significant higher at sandy soils, this difference was not automatically caused by soil conditions. This contrast could be caused by the mean age of both categories. Sampling plots on sandy soil had a mean age of 3.3 years, while the mean sampling plot on clay soils was 2.8 years old. Although this seems like a slight difference, it could be a reason for enhanced AGC stock on sand soils. To prevent the analysis for being influenced by age, a comparison per age category (in number of growing season) could be made. However, the dataset was too small to do so. Furthermore, age is just one of many variables that could influence the comparison between soil texture classes. Former land use could not have had an influence, because food forests with FLU *forest* were excluded from this comparison and no significant difference was found between former arable lands and grasslands. But, for example, environmental conditions could have influenced the effect of soil texture on AGC stock. For example, the included food forests had a variation in planting design, management and abiotic circumstances. Some of the food forests had extreme high moisture values during winters (*e.g.* Mijn Stadstuin) caused by the location of the food forest. Others were suffering with common high winds (*e.g.* Kreilerwoud), while others were well sheltered. A relatively low growth rate (and AGC accumulation) in Mijn Stadstuin and Kreilerwoud might not (or only partially) be caused by their soil texture but caused by their unfavourable environmental conditions. Therefore, it is hard to prove the effect of FLU on carbon stock development. Since long periods of drought occurred more frequently in the Netherlands over the last couple of years (Philip *et al.,* 2020), and food forests on sandy soils seemed to suffer more with these drought than food forests on clay soils (pers. comm.), the hypothesis composed in the beginning of this study still remains.

The structural composition of food forests was also declaring variety in aboveground carbon stock. As hypothesised, dispersed food forests had a larger AGC stock than food forests with alleys, due to positive interspecific interactions (Morin *et al.,* 2011; Palandrani, Battipaglia & Alberti, 2020). These differences could not be declared by a difference in soil texture or former land use since these variables were constant in the analysed plots (this was one of the major benefits of applying zones). Age could have influenced the difference between both compositions, since the dispersed zones were older than the zones with alleys in two food forests (Ketelbroek and Benthuizen). However, even when these were both excluded, a significant difference between *dispersed* and *alleys* was found. The main disadvantage of this analysis is the small number of food forests that were included (five) and that they were all relatively young (at most 3 years old). The impact of planting is therefore expected to be relatively large. Dispersed food forests might be planted more intensively than alleys since alleys were designed to maximise yield and harvest in future. Whether the difference between these two structural compositions still occur in more mature food forests is questionable. The hypothesis of a positive effect of interspecific interactions in dispersed food forests could not be confirmed based on the data obtained in this study, since these effects are not expected to be visible within one or two years (Jose, Gillespie & Pallardy, 2004; Ong *et al.,* 1991; Asthon, 1999). Planting statistics could be analysed for these specific food forests, in order to exclude the effect of planting on this difference. Although dispersed food forests could provide significant yields as well, large-scale harvest is more effective in a food forest with an alley-design (Chaturvedi, 1992; Ferguson & Lovell, 2014). If food forest with alleys are providing as much carbon stock as dispersed food forests, this might enlarge the interest in the creation of new food forests and the capacity of food forests to fungate as affordable carbon storage projects (Riolo, 2019; Opiniepanel, 2019). Even if carbon storage in an alley-design does not equal the storage in a dispersed food forest, one could consider the trade-off between harvest-friendliness and carbon storage. Therefore, it is necessary to determine whether there is a difference in carbon stock between both compositions and what this difference exactly is.

5.3. Influence of soil variables on AGC stock

Out of all soil variables, the amount of soil organic matter (SOM) was most strictly correlating with aboveground carbon stock. However, the regression between SOM and AGC stock was not significant, neither was the correlation between both. The hypothesis of an increase in moisture and nutrient availability in SOM rich soils (Billings, 2006; Craswell & Lefroy, 2001) was confirmed concerning food forests included in this study (De Groot, 2020). This could be declared by two things. First, food forests with a large AGC stock developed a large amount of SOM in their soils. Second, food forests with a large amount of SOM were able to accumulate carbon more quickly than SOM poor food forests, due to positive influence of SOM on plant growth (Craswell & Lefroy, 2001; Bot & Benites, 2005). In fact, both aspects are expected to be true, but the second aspect has not been proven in this research. Food forests with a different initial SOM content should be monitored over several years to determine whether carbon stock is accumulating quicker in presence of large amounts of SOM indeed. For example, food forests Eemvallei-Zuid and Schijndel-Hardekamp are suitable for this analysis. Food forest Eemvallei-Zuid is located on a soil with a relatively high initial SOM content of 4.1, while Schijndel-Hardekamp is located on a soil with a significant lower amount of SOM (2.7, in essence the second lowest of all food forests). Both food forests are young (1 and 2 years old respectively) and relatively large-scale, making them comparable and appropriate for extensive measurements. One should consider that the soils of both food forests do not have the same texture. Food forest Benthuizen could be used as alternative for Eemvallei-Zuid, to increase the difference in SOM content (from 1.4 to 7.3). However, Benthuizen is way smaller than Schijndel-Hardekamp and besides a former peatland with high water level. This larger variety on multiple variables would increase the difficulty of data interpretation, making the comparison between Eemvallei-Zuid en Schijndel-Hardekamp more favourable. In conclusion, one could analyse the differences in carbon accumulation between two food forests, but there are always differences concerning other variables that should be borne in mind. An expansion of the total dataset by the addition of more food forests should make the analysis of the relation between SOM and AGC stock more accurate.

The large variation in initial SOM conditions between the food forests reinforces the choice to compare Dutch food forests based on their aboveground carbon stock only. As described before, the effects of land use on soil conditions could remain visible for decades (Bissett, 2011; Callaham, 2006). An assessment of the total carbon stock of a food forest has therefore a high chance of being influenced by the former land use. For example, food forest Benthuizen is relatively young, but has a large soil carbon stock. In fact, one should include the additional soil organic carbon content, from the moment the land was converted to a food forest (at least for the former arable lands and grasslands). This matter requires an extensive assessment, which could not be performed in this study. Furthermore, soil organic carbon is much more volatile than carbon stored in biomass (McBride *et al.,* 2020; Buchanan & King, 1992). Lastly, aboveground carbon stock accumulates quicker than a soil carbon stock. Both aspects speak in favour of the use of aboveground carbon stocks to compare (the relatively young) Dutch food forests in a standardised and robust way.

5.4. Microclimate of food forests

The results concerning microclimate of food forests were partly identical to what was hypothesised (Konarska *et al.,* 2016; Van Noordwijk *et al.,* 2014; Gosme *et al.,* 2016). On average, food forests had a cooling and moisturizing effect on the air. This effect is comparable to the effect of natural forests (Midrexler, Zhao & Running, 2011). Hardly any food forests without an effect on temperature and humidity have been found. It is remarkable that air is already cooler and moisture in the youngest food forests. Most of the young food forests still had a relatively small carbon stock and were characterised by large herbaceous vegetation. The observed difference in humidity and temperature in these young food forests implicates that the presence of herbaceous vegetation only affected the microclimate already. This is found in other studies on herbaceous vegetation as well (Kaufmann *et al.,* 2003; Simonin *et al.,* 2014). But as described in the hypotheses, forests are expected to have a larger effect on temperature and humidity than grassland vegetation (Teuling *et al.,*2010). In essence, the relative differences between inside and outside measurements were expected to increase over aboveground carbon stock. This effect has not been observed in this research. This could be caused by the absence of measurements in food forests with a large carbon stock. In food forest Voedselrijk and Sualmana (second and third highest AGC stock of all food forests respectively), the outside measurements were also executed in a forest. Even though it was outside the food forest, these measurements did not comply with the protocol of measuring away from wooded areas. The other food forest with a large carbon stock (De Stomp) did not show a relatively large difference in temperature and humidity either. But these measurements could be affected by the extremely cold and wet weather conditions that day, which were exceptions to generally warm and dry measurement days. Although the differences in temperature and humidity were transformed to relative values, these extreme conditions could have been influencing. In short, the presence or absence of a regression between microclimate conditions and AGC stock should be confirmed in further, more extended research. More accurate measurements in food forests with large aboveground carbon stocks are necessary to be included in those studies, even if that means that outside temperature and humidity will not be monitored just outside the food forest. The hypothesis of an increasing effect of food forests on microclimate conditions as aboveground carbon stock grows, remains intact.

5.5. Elements of aboveground carbon stock calculations

To relation between the three characteristics canopy closure, maximum height and basal area and aboveground carbon stock was analysed to determine whether these characteristics are appropriate to determine AGC stock in food forests. A significant regression between canopy closure and aboveground carbon stock was found, although this regression was not linear but logistic. At low AGC stock levels canopy closure was increasing strongly, but at higher AGC stock levels the regression flattens. Canopy closure might therefore be a good explanatory variable for young food forests, but the variation is too large to predict carbon stock in more mature forests. The turning point seemed to be located at an AGC stock of 15 Mg ha^{-1}, after which canopy did not increase any further when AGC stock increased. A food forest without any initial AGC stock (*i.e.* food forests on former arable lands and grasslands) are expected to reach this carbon stock after 7-10 years, based on the AGC stocks found in this study. For example, a canopy closure of 80% could be related to carbon stocks of 40 to 400 Mg ha^{-1} in this study. Within the first years of food forest development (especially on arable lands and grasslands), canopy closure might be useful as quick method to estimate carbon stock. But even in immature food forests, canopy closure is not very accurate. An AGC stock of 4.4 Mg ha^{-1} was found in a plot with 0% canopy closure, while another plot had a canopy closure of 34% and a AGC stock of 0.9 Mg ha^{-1}. These large variations could be caused by the inaccuracy of measurements, but these are found in other studies as well (Valverde & Silvertown, 1997; Heynen & Lindsey, 2003) and should be considered when using canopy closure as estimator of AGC stock. When these relatively large confidence intervals are considered, canopy closure can be used as quick, simple estimator of AGC stock. Furthermore, it might be interesting to keep measuring canopy closure for other purposes, such as light availability for herbaceous vegetation (Parent & Messier; 1996; Warren *et al.,* 2013).

In contradiction to canopy closure, the maximum height of the tree layer was describing variation in aboveground carbon stock surprisingly good. Compared to the regression of canopy closure, the regression of maximum height over AGC stock had less deviation and a higher explanatory capacity (based on the adjusted R^2 of 0.87 and 0.66 respectively). Furthermore, the regression seemed much more monotonous, following an exponential curve, which was transformed into a perfect linear relationship after logarithmic transformation of carbon stock. Based on existing literature, this strong relation was not expected (Vieilledent *et al.,* 2012; Mensah, Veldtman & Seifert, 2017). Since food forests are very dissimilar, concerning *e.g.* former land use, planting strategy, species composition and planting density, a large variety of AGC stock at a certain maximum tree height was expected (Levy *et al.,* 2004; Calvo-Alvarado, McDowell & Waring, 2008). The total basal area of the plot was also significantly correlating with AGC stock, even without a logarithmic transformation, with an adjusted R^2 of 0.84. Both maximum height of the tree layer and basal area seemed to be an accurate estimator of AGC stock. However, a model including both variables was explaining significantly more variation in AGC stock, with an adj. R^2 of 0.99. This result is not surprising, since both tree carbon stock measurements include both variables in their allometric equations (Verra; UNFCCC, 2013). Furthermore, Verschuyl *et al.* (2018) concluded that shrub carbon stock could be best predicted by an allometric equation including both basal area and maximum height. Nevertheless, the strong and significant regressions between AGC stock and both variables individually are interesting, as this might insinuate that AGC stock could be estimated quite accurately based on one variable (food forest height or basal area) only. The appropriateness of both variables as explanatory variables could be examined in future. For example, a more extended and precise data sampling technique of maximum tree height could be determined. Furthermore, Multiple allometric equations with increasing complexity could be compared with each other (instead of only comparing *glm's*). If field work methodologies could be made less time-consuming due to the discovery of an accurate, simple model of estimating carbon stock, more food forests could be included in one study and/or fieldwork could be focussed on more in-depth analyses. In future, one could distinguish a complex sampling technique including basal area, canopy closure and maximum height to determine AGC stock precisely and a simple sampling technique including only basal area or maximum height to estimate AGC stock roughly.

In order to determine the usefulness of these measurements, an analysis of shrub AGC stock has been made. The relative amount of carbon stored in shrubs was negatively correlating with increasing carbon stock. However, the variation in AGC stock of shrubs predicted by this correlation was very low (adj. $R^2 = 0.09$) making the correlation less meaningful. Furthermore, this negative relation was mainly caused by the former production forests (Voedselrijk and De Stomp). These two food forests had the largest AGC stock, but a relatively small shrub carbon stock. Since hardly any shrubs were planted in these former production forests yet, the limited amount of carbon stored in shrubs in these food forests is not surprisingly. Excluding the former forests from this analysis gave other insights. In food forest Ketelbroek, which is the second oldest food forest planted with FLU *arable land* or *grassland*, 8.5% of AGC was stored in shrubs (at an age of 9 years on average). Assessments on the proportion of carbon stored in shrubs in natural forests are limited, but multiple studies estimated that the AGC stocks of shrubs are a maximum of 3% (Ullah & Al-Amin, 2012, Janssens *et al.,* 1999). A study in Nepal found a relative carbon stock in shrubs of 10%, however climatological circumstances were different from the situation in The Netherlands (Dangal, Das & Paudel, 2017). The relative carbon stock of 8.5% in food forest Ketelbroek therefore seemed to be relatively high, compared to natural forests, making the inclusion of shrub measurements in assessments on carbon stocks of food forests useful. It was not possible to determine a fixed percentage of shrub AGC stock, based on the data obtained in this study. The correlation that described this fixed ratio only had a coefficient of 0.55 and older food forests were underrepresented. In conclusion, the hypothesis that shrubs contributed significantly to the AGC stock of food forests could neither be rejected nor confirmed, and it would be sensible to continue shrub measurements in order to determine potential fixed percentages in future.

5.6. Food forests versus natural forests

On average, AGC stock was significantly lower in food forests than in the reference forests included in this research. It is assumable that this difference is mainly caused by the higher mean age of reference forests (Nabuurs & Mohren, 1993; Daamen, 2008; Schelhaas & Clerkx, 2015). Therefore, these reference forests are not useful in carbon stock comparisons (although they could explain unexpected data points in the trends of carbon accumulation in future). It might be interesting to expand the study on reference forests with secondary growing forests, for example after forest fires and/or reforestation projects. The exact age of these forests is known, and the increase in carbon stock can be monitored within the first years of succession. When using reforestation projects as reference, one should take the purpose of maximising carbon storage (Face the Future, Dybala *et al.,* 2019) into account. Production forests had the same objective since these were designed to produce timber. Food forests in former production forests (Voedselrijk, De Stomp) had indeed a higher initial carbon stock than the reference forests. Former production forests that have had the same objective (Voedselrijk en De Stomp) have shown a significantly higher AGC stock than the natural forests measured in this study. Carbon stock in these reforestation areas might therefore accumulate quicker than can be expected in food forests (and natural forests), making secondary forests most appropriate as natural forest reference.

When comparing secondary temperate forests of 20 years old (Pregitzer & Euskirchen 2004; Rooduijn *et al.,* unpublished) to the oldest food forest in this research (Sualmana, 22 years old), natural forests had a higher carbon stock, with a difference of 21%. We hypothesised an equal carbon stock in food forests and in natural forests, like found in tropical agroforestry systems (61.5 and 60.9 Mg ha^{-1} on average respectively (Rooduijn *et al.*, unpublished). As mentioned before, there are arguments to suspect an unrepresentatively low carbon stock in Sualmana. First, the food forest is located at a dry, sandy soil, with a low water availability. Second, nutrients have been removed from the food forest, impoverishing the nutrient availability in the soil. Both processes were negatively influencing tree growth and therefore aboveground carbon accumulation (Basset, 1994; Sullivan *et al.,* 2015). Based on this study, we cannot reject the hypothesis of a similar carbon stock in food forests and natural forests, since this would completely be based on one food forest only. Whether other food forests transcend the carbon stock of Sualmana (37 Mg ha^{-1} after 22 years) and reaches the carbon stock of temperate natural forests (45 Mg ha^{-1} after 20 years) should become clear in the coming years. But it is important to consider that every food forest is different and that carbon stocks of one or two single cases could not automatically be representative for all food forests. For example, Ketelbroek has a relatively high plant diversity and is not designed to maximise harvest. The expected carbon stock after 20 years is therefore higher in Ketelbroek than in Sualmana, but not automatically more representative for food forests in general. The carbon stock of food forests at a certain age should therefore be determined by a weighted average of standardised results, using single cases to explain variation in AGC stocks between food forests.

Food forests with FLU *forest* have been compared to natural forests as well. This selection of food forests also had a significant lower AGC stock the reference forests. This could be declared by the fact that 2 out of 6 food forests in FLU category *forest* had a relatively low initial carbon stock. These food forests were converted from an orchard and a tree nursery. The former production forests had a similar carbon stock as the reference forests. This is remarkable since production forests were expected to have maximised carbon stock (and the natural forest did not). Because the reference forests were measured very extensive (using only one sampling plot each), the results are assumed to be not very reliable. For example, a natural forest typically has very dense and less dense zones, *i.e.* it is quite heterogeneous (Zenner & Hibbs, 2000). Moreover, the natural edges of a forest contain less ABG carbon than the inner parts of the forest (Chaplin-Kramer *et al.,* 2015; Remy *et al.,* 2016). Because the sampling plots were generally located in a dense area and away from forest edges, the calculated carbon stock of reference forests could be higher than it actually was. This could declare the fact that carbon stocks measured in natural forests in this research were relatively high compared to other studies (Nabuurs & Mohren, 1993; Sikkema & Nabuurs, 1994). However, as mentioned before, it is hard to determine to what extent a Dutch forest is natural. It is assumable that the reference forests have been planted as production forests as well (Veenman, Lieferink & Arts, 2009), making the carbon stock of reference forests comparable to the carbon stock of productions forests. This corresponds to the data obtained in this study (mean AGC stock of 197 ± 78 Mg ha^{-1} in reference forests and 140 ± 110 Mg ha^{-1} in former production forests).

When comparing the AGC stock of food forests to other studies, differences in methodology should be considered. The methodology in this research is mainly based on verified methodologies used all over the world (UNFCCC, 2013). However, these standards are not directly copied. The major modification was the inclusion of shrubs and small trees with a diameter at breast height of less than 5 centimetres. This adjustment increased the accuracy of carbon stock estimates but could have led to a higher AGC stock than studies that did not include these plants. Although the percentage of carbon stored in shrubs and small trees is expected to be relatively low, it could be one of the reasons that food forest Sualmana had a higher carbon stock than the Agroforestry Research Trust food forest in Devon (Lehmann *et al.,* 2019; Schafer, Lysák & Henriksen, 2019). One could assume that the carbon stock calculations (*i.e.* the allometric equations and its assumptions) in other studies was similar in general, as the VCS was mainly followed. An extended analysis on the exact equations and implementations used in other studies on (food) forest carbon stocks could confirm the presumption of carbon stocks being comparable with each other. Furthermore, it might be useful to contact the managers and researchers involved in the Agroforestry Research Trust food forest to adjust both methodologies to one another.

The aboveground carbon stock of the oldest food forest included in this research was significantly lower than the mean aboveground carbon stock of tropical home gardens (Rooduijn *et al.,* unpublished data). Since temperate systems store more carbon belowground (relatively to their aboveground carbon storage (Birch, 1999), this result is not surprising. The fact that the AGC stock of temperate food forests is lower than tropical home gardens and agroforestry systems does therefore not mean that the total amount of carbon stored by the temperate food forest is lower as well. This food forest focussed on the variation in carbon stocks between Dutch food forests, which can be analysed based on AGC stock only. An estimate of total carbon stocks of food forests is necessary to facilitate a fair comparison between temperate and tropic systems.

5.7. Limitations of this study

The main limitation of this study is the underrepresentation of older food forests. As we are investigating a whole new concept of agriculture, we are just in the baseline stage of examining food forests. All analyses on the development of carbon stock over food forest age are suffering with the many variables that could have influenced this development. But one of the major goals of the national monitoring program food forests (NMPF) was to set up large-scale, standardised database. If we continue down this road, the actual increase in AGC stock in the same sampling plot could be determined. The results described and discussed in this study can therefore be seen as baseline investigation and can be used as a guideline for further research.

The objective of the NMPF to standardise and enlarge datasets on food forests, also has a downside. As hypothesised in advance, most food forests are very heterogeneous. By including so many food forests in this study, the sampling effort per food forest was limited. Several analyses, such as the exact effect of two soil texture classes within one food forest, were impracticable due to a lack of sufficient data. An expansion of the standard measurements with extra sampling points for a selection of interesting food forests would have made these analyses more expressive. In the coming years, it might be useful to make time for these additional measurements.

Another aspect that was not analysed in this study is the presence of temporary trees. Some of the food forests planted temporary, fast-growing trees, in order to create favourable conditions for the more slowly growing crops. For example, these temporary trees provided shade and increased the moisture retain capacity of the soil. When the intended canopy trees have grown enough to survive itself, the temporary trees will be removed from the food forest. This could lead to an unexpected decrease in aboveground carbon stock. The presence of temporary trees could insinuate a more rapid carbon accumulation. Since these individuals were not present in all food forests, differences in carbon stock accumulation rates might partly be caused and explained by the presence of these temporary trees. This aspect could be one of the driving factors of the unexplained variety in AGC stock of food forests.

At last, the implementation of zone selection in the methodology was both beneficial and limiting this study. In some situation, such as the examples mentioned in the previous paragraph, the implementation of zones provided useful information about the effects of variables on AGC stock. However, in some situations, the selected zones were hardly applicable in comparisons, for example when zones were distinguished based on multiple variables. Food forest Benthuizen was divided in two zones based on both age and structural composition. Since these two variables were both expected to affect aboveground carbon stock, these zones did not provide any specific information on the actual effect of age nor structural composition. Moreover, some of the zones were clearly distinguishable, but these differences were difficult to express in specific variables (for example in food forests De Overtuin and Voedselrijk). In these situations, the implementation of zones increased sampling effort, but did not increase the usefulness of the data. One could argue that it is more efficient to use a fixed sampling effort in every food forest, as been done in the study on soil conditions. In this way, time could be made for additional measurements in multiple zones, in food forests where this is expected to be appropriate only. Furthermore, the almost impossible objective to standardise zone selection could be released this way. From another perspective, it might be too early to classify some of the zones as useless. When comparing measurements on aboveground carbon stock in the same sampling plots each year, differences in accumulation could be compared between zones. Even the complex zones in De Overtuin and Voedselrijk could show differences in carbon accumulation rate. This could (unexpectedly) lead to new insights and further research can be based on these differences. Therefore, the implementation of zones as performed in this study could still be useful in all food forests, but the selection of specific food forests for additional, expanded measurements is a good idea anyway.

5.8. Recommendations for further research

As the results of this study were discussed, more and more interesting analyses came up, that could not be included in this research anymore due to time limitations. Furthermore, this study was limited in several ways, some of which could have been prevented by adjustments on methodology or research approach. These things combined have led to the following list of recommendations for further research. Note that many of these suggestions were explained in *Discussion*.

Validation of the results of this study

- Analyse the accumulation of aboveground carbon stock over time within the same sampling plots, to determine the exact development of AGC stock in food forests.
- Examine the effect of increasing aboveground carbon stock on the microclimate of food forests, including temperature and humidity analyses. A new, appropriate methodology for food forests located within natural forests needs to be made.
- Analyse the planting statistics of food forests, in order to determine whether differences in aboveground carbon stock between *alleys* and *dispersed* food forests were caused by interspecific relations or planting techniques.

Extension of analyses of this study

- Determine the total carbon stock of the selected food forests, including aboveground carbon stock and amount of carbon stored in the soil. Ideally, measurements on food forest litter will be included as well.
- Expand the comparison between aboveground carbon stock and soil variables. In this study, only pH, moisture availability, cation exchange capacity, total amount of nitrogen and soil organic matter were compared to AGC stock, but dozens of other variables are available.
- Expand the sampling effort in a selected number of food forests, in order to determine the exact effect of one specific variable on aboveground carbon stock. Food forest Roggebotstaete is suggested for the effect of soil texture; food forest Lekker Landgoed seems to be most suitable for analyses on the effect of age; an ideal location to analyse the effect of former land use is not available yet. Note that case studies are only desirable to support standardised and large-scale studies, not to replace them.
- Analyse the differences in aboveground carbon stock development between zones in food forest De Overtuin en Voedselrijk. The differences between these zones were based on (amount other things) permaculture management zones and future differences in structure. Due to its complexity, the zones were not compared to each other in this study, but an extended analysis could lead to new, important insights.

Expansion of the dataset

- Include more extended measurements on natural forests as reference for food forest aboveground carbon stock. Ideally, sampling plots in secondary developing forests and reforestation projects were added to the current selection of reference forests.
- Include more data of relatively old temperate food forests and synchronise methodology with that of older food forests in other temperate areas, to validate the aboveground carbon stock of the oldest food forests in the Netherlands.

Improvement of the methodology

- Examine whether the maximum height of the tree layer or basal area of food forests is an appropriate variable for low-threshold aboveground carbon stock estimates.
- Examine whether there is a fixed percentage of carbon stored in aboveground biomass of shrubs, in order to make the measurements less time-consuming. At least several years of measurements are expected to be necessary before this assessment could be made.

6. Conclusions

The accumulation of aboveground carbon over food forest age seemed to follow a sigmoid curve. However, the accuracy of the saturation of this curve is uncertain. The development of carbon stock in the oldest food forests has to be analysed to confirm the sigmoid curve of carbon accumulation in food forests. AGC stock of food forests was significantly different between categories of structural composition and former land use. *Dispersed* food forests had a higher AGC stock than food forests with *alleys*. Whether positive interspecific interactions have played a role on this difference is questionable. Former forests had a larger AGC stock than former grasslands and arable lands. A decrease of AGC stock over food forest age is found in these former forests, possibly caused by thinning of the food forests or the large variety within this FLU category. An increase of carbon stock is expected after several years in food forests with FLU *forest* as well. Whether carbon accumulation is higher in former forests than on former arable lands and grasslands could not be concluded. Soil conditions seemed to have little effect on carbon accumulation. No significant effect of soil texture class on food forest aboveground carbon stock was found in this study, neither a significant correlation between AGC stock and one of the five soil variables (acidity, amount of SOM, amount of nitrogen, nutrient-, and water availability). An analysis of AGC accumulation over time within the same plots is necessary to confirm the absence of any effect of soil conditions on carbon accumulation. This analysis could also confirm the effect of structural composition and former land use on AGC stock.

Food forests had a significant effect on the microclimate. A significant higher temperature and lower humidity of the air was found inside food forests compared to the outside conditions. No significant regressions of these differences in in temperature and humidity over food forest age or AGC stock were found. However, due to a lack of relatively old food forests and food forests with a relatively large AGC stock in these measurements, the presence nor absence of this regression could be determined. Even young food forests without a large AGC stock had a cooling and moisturising effect on the microclimate.

The mean AGC stock of food forests was significantly lower than that of reference forests. The accuracy of measurements on reference forests is debatable. Food forests included in this study seemed not to reach the same amount of carbon stock found in temperate natural forests. However, these conclusions are solely based on one single, relatively old food forest, and are therefore uncertain. A comparison between food forests and reforestation projects and secondary forests should be useful. Food forests did not store as much carbon as tropical agroforestry systems in their aboveground biomass. However, since temperate forests have a higher relative belowground carbon stock than tropical forests, this does not automatically mean that the total carbon stock of temperate food forests is lower than its tropical equivalents. Further research is necessary to describe this relation between temperate and tropic systems.

Both basal area and maximum height were explaining a large proportion of the variation in AGC stock (84 and 87% relatively). A combination of both variables was explaining almost all variation in AGC stock (99%). Canopy closure was explaining less variation (66%) and an addition of canopy closure to the model did not increase the explaining capacity (99%). Basal area and maximum height seemed to be useful variables to estimate aboveground carbon stock, although it is necessary to measure both in order to determine AGC stock more precisely. A fixed percentage of AGC stored in shrubs could not be determined and therefore it is useful to continue shrub carbon stock measurements.

References

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. Encyclopedia of research design, 3, 583-585.

Akaike, H. (1974), "A new look at the statistical model identification", IEEE Transactions on Automatic Control, 19 (6): 716–723, doi:10.1109/TAC.1974.1100705, MR 0423716.

Asase, A., & Tetteh, D. A. (2010). The role of complex agroforestry systems in the conservation of forest tree diversity and structure in southeastern Ghana. Agroforestry systems, 79(3), 355-368.

Ashton, P. S. (1999). Ecological theory of diversity and its application to mixed species plantation systems. The silvicultural basis for agroforestry systems. CRC Press, Boca Raton, 61-77.

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O., Swartz, B., Quental, T. B., ... & Mersey, B. (2011). Has the Earth's sixth mass extinction already arrived?. Nature, 471(7336), 51-57.

Barton, K., & Barton, M. K. (2015). Package 'MuMIn'. Version, 1, 18.

Bassett, J. R. (1964). Tree growth as affected by soil moisture availability. Soil Science Society of America Journal, 28(3), 436-438.

Baty, F., Ritz, C., & Baty, M. F. (2015). Package 'nlstools'. Tools for Nonlinear Regression Analysis.

Besnard *et al.* 2018 (Quantifying the effect of forest age in annual net forest carbon balance)

Biesmeijer, J. C., Roberts, S. P., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., ... & Settele, J. (2006). Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. Science, 313(5785), 351-354.

Billings, S. A. (2006). Soil organic matter dynamics and land use change at a grassland/forest ecotone. Soil Biology and Biochemistry, 38(9), 2934-2943.

Birch, C. P. (1999). A new generalized logistic sigmoid growth equation compared with the Richards growth equation. Annals of Botany, 83(6), 713-723.

Bissett, Andrew, Alan E. Richardson, Geoff Baker, and Peter H. Thrall. "Long-term land use effects on soil microbial community structure and function." Applied Soil Ecology 51 (2011): 66-78.

Björklund, J., Araya, H., Edwards, S., Goncalves, A., Höök, K., Lundberg, J., & Medina, C. (2012). Ecosystem-Based Agriculture Combining Production and Conservation—A Viable Way to Feed the World in the Long Term?. Journal of Sustainable Agriculture, 36(7), 824-855.

Bot, A., & Benites, J. (2005). The importance of soil organic matter: Key to drought-resistant soil and sustained food production (No. 80). Food & Agriculture Organisation.

Bourdeau, P. F. (1953). A Test of Random Versus Systematic Ecological Sampling. Ecology, 34(3), 499-512.

Brockerhoff, E. G., Barbaro, L., Castagneyrol, B., Forrester, D. I., Gardiner, B., González-Olabarria, J. R., ... & Thompson, I. D. (2017). Forest biodiversity, ecosystem functioning and the provision of ecosystem services.

Brooks, R. T., Nislow, K. H., Lowe, W. H., Wilson, M. K., & King, D. I. (2012). Forest succession and terrestrial–aquatic biodiversity in small forested watersheds: a review of principles, relationships and implications for management. Forestry, 85(3), 315-328.

Brooks, T. M., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., Rylands, A. B., Konstant, W. R., ... & Hilton-Taylor, C. (2002). Habitat loss and extinction in the hotspots of biodiversity. Conservation biology, 16(4), 909-923.

Brown, R. H. (1984). Growth of the green plant. Physiological basis of crop growth and development, 153-174.

Brüning, L. Z., Krieger, M., Meneses-Pelayo, E., Eisenhauer, N., Pinilla, M. P. R., Reu, B., & Ernst, R. (2018). Land-use heterogeneity by small-scale agriculture promotes amphibian diversity in montane agroforestry systems of northeast Colombia. Agriculture, ecosystems & environment, 264, 15-23.

Buchanan, M., & King, L. D. (1992). Seasonal fluctuations in soil microbial biomass carbon, phosphorus, and activity in no-till and reduced-chemical-input maize agroecosystems. *Biology and Fertility of Soils*, *13*(4), 211-217.

Callaham Jr, M. A., D. D. Richter Jr, D. C. Coleman, and M. Hofmockel. "Long-term land-use effects on soil invertebrate communities in Southern Piedmont soils, USA." European Journal of Soil Biology 42 (2006): S150-S156.

Calvo-Alvarado, J. C., McDowell, N. G., & Waring, R. H. (2008). Allometric relationships predicting foliar biomass and leaf area: sapwood area ratio from tree height in five Costa Rican rain forest species. Tree physiology, 28(11), 1601-1608.

Cao, T., Valsta, L., & Mäkelä, A. (2010). A comparison of carbon assessment methods for optimizing timber production and carbon sequestration in Scots pine stands. Forest ecology and management, 260(10), 1726-1734.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., ... & Kinzig, A. P. (2012). Biodiversity loss and its impact on humanity. Nature, 486(7401), 59-67.

Cavoski, I., Caboni, P., & Miano, T. (2011). Natural pesticides and future perspectives. Pesticides in the modern world-pesticides use and management, 169-190.

Chaplin-Kramer, R., Ramler, I., Sharp, R., Haddad, N. M., Gerber, J. S., West, P. C., ... & Mueller, C. (2015). Degradation in carbon stocks near tropical forest edges. Nature communications, 6(1), 1-6.

Chappell, M. J., & LaValle, L. A. (2011). Food security and biodiversity: can we have both? An agroecological analysis. Agriculture and Human Values, 28(1), 3-26.

Chaturvedi, A. N. (1992). Optimum rotation of harvest for poplars in farmlands under agroforestry. Indian forester, 118(2), 81-88.

Ciais, Philippe, Mart-Jan Schelhaas, Sönke Zaehle, S. L. Piao, Alessandro Cescatti, Jari Liski, Sebastiaan Luyssaert *et al*. "Carbon accumulation in European forests." Nature Geoscience 1, no. 7 (2008): 425-429.

CL Compendium voor de Leefomgeving. (2018). Fauna van het agrarisch gebied, 1990-2018 | Compendium voor de Leefomgeving. Retrieved October 26, 2020, from https://www.clo.nl/indicatoren/nl1580-trend-fauna-agrarisch?ond=20877.

Collinge, S. K. (2000). Effects of grassland fragmentation on insect species loss, colonization, and movement patterns. Ecology, 81(8), 2211-2226.

Conservation International. Retrieved 2020, from https://www.conservation.org/blog/massive-reforestation-effort-puts-down-roots-in-brazilian-amazon/

Coomes, D. A., & Grubb, P. J. (2000). Impacts of root competition in forests and woodlands: a theoretical framework and review of experiments. Ecological monographs, 70(2), 171-207.

Craswell, E. T., & Lefroy, R. D. B. (2001). The role and function of organic matter in tropical soils. In Managing Organic Matter in Tropical Soils: Scope and Limitations (pp. 7-18). Springer, Dordrecht.

Craswell, E. T., & Lefroy, R. D. B. (2001). The role and function of organic matter in tropical soils. In Managing Organic Matter in Tropical Soils: Scope and Limitations (pp. 7-18). Springer, Dordrecht.

Daamen, W. P. (2008). Kaart van de oudste bossen in Nederland: Kansen op hot spots voor biodiversiteit (No. 121). Wettelijke Onderzoekstaken Natuur & Milieu.

Dalen, L.S. (2017) Forests play important part in cooling down local climate. A new study shows how important forests are in keeping much of the planet's surface cool.

Dangal, S. P., Das, A. K., & Paudel, S. K. (2017). Effectiveness of management interventions on forest carbon stock in planted forests in Nepal. Journal of environmental management, 196, 511-517.

De Groot, E.M. (2020). Exploring soil restoration potential in Dutch food forests - a baseline study of soil ecological performance.

De Vos, J. M., Joppa, L. N., Gittleman, J. L., Stephens, P. R., & Pimm, S. L. (2015). Estimating the normal background rate of species extinction. Conservation biology, 29(2), 452-462.

Dewar, R. C. (1990). A model of carbon storage in forests and forest products. Tree physiology, 6(4), 417-428.

Díaz, S., Fargione, J., Chapin III, F. S., & Tilman, D. (2006). Biodiversity loss threatens human well-being. PLoS Biol, 4(8), e277.

Dimoudi, A., & Nikolopoulou, M. (2003). Vegetation in the urban environment: microclimatic analysis and benefits. Energy and buildings, 35(1), 69-76.

Dutch Ministry of Nature. (2020). NL Pollinator Strategy. Retrieved from https://www.government.nl.

Dybala, K. E., Steger, K., Walsh, R. G., Smart, D. R., Gardali, T., & Seavy, N. E. (2019). Optimizing carbon storage and biodiversity co-benefits in reforested riparian zones. Journal of Applied Ecology, 56(2), 343-353.

Evrendilek, F., Celik, I., & Kilic, S. (2004). Changes in soil organic carbon and other physical soil properties along adjacent Mediterranean forest, grassland, and cropland ecosystems in Turkey. Journal of arid environments, 59(4), 743-752.

Face the Future, Retrieved 2020, from https://facethefuture.com/projects/nederland.

Fearnside, P. M., & Laurance, W. F. (2004). Tropical deforestation and greenhouse-gas emissions. Ecological Applications, 14(4), 982-986.

Ferguson, R. S., & Lovell, S. T. (2013). Permaculture for agroecology: design, movement, practice, and worldview. A review.

Ferguson, R. S., & Lovell, S. T. (2014). Permaculture for agroecology: design, movement, practice, and worldview. A review. Agronomy for Sustainable Development, 34(2), 251-274.

Fischer, J., Abson, D. J., Butsic, V., Chappell, M. J., Ekroos, J., Hanspach, J., ... & von Wehrden, H. (2014). Land sparing versus land sharing: moving forward. Conservation Letters, 7(3), 149-157.

Foley, Jonathan A., Ruth DeFries, Gregory P. Asner, Carol Barford, Gordon Bonan, Stephen R. Carpenter, F. Stuart Chapin *et al.* "Global consequences of land use." science 309, no. 5734 (2005): 570-574.

Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., ... & Graves, S. (2016). car: An R Companion to Applied Regression. R package version 3.2-0.

Geiger *et al* 2010 Geiger, F., Bengtsson, J., Berendse, F., Weisser, W. W., Emmerson, M., Morales, M. B., ... & Eggers, S. (2010). Persistent negative effects of pesticides on biodiversity and biological control potential on European farmland. Basic and Applied Ecology, 11(2), 97-105.

Girvan, M. S., Bullimore, J., Ball, A. S., Pretty, J. N., & Osborn, A. M. (2004). Responses of active bacterial and fungal communities in soils under winter wheat to different fertilizer and pesticide regimens. Applied and environmental microbiology, 70(5), 2692-2701.

Gosme, M., Dufour, L., Aguirre, H. D. I., & Dupraz, C. (2016, May). Microclimatic effect of agroforestry on diurnal temperature cycle. In 3. European Agroforestry Conference (EURAF 2016) (pp. 466-p).

Green Deal Food Forests (2019). Retrieved 2020, from https://www.greandeelvoedselbossen.nl/green-deal-voedselbossen/.

Green, R. E., Cornell, S. J., Scharlemann, J. P., & Balmford, A. (2005). Farming and the fate of wild nature. science, 307(5709), 550-555.

Groenendyk, D. G., Ferré, T. P., Thorp, K. R., & Rice, A. K. (2015). Hydrologic-process-based soil texture classifications for improved visualization of landscape function. PLoS One, 10(6), e0131299.

Gross, G. (2012). Effects of different vegetation on temperature in an urban building environment. Micro-scale numerical experiments. Meteorologische zeitschrift, 21(4), 399-412.

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020). dplyr: A Grammar of Data Manipulation. R package version 1.0.2. https://CRAN.R-project.org/package=dplyr.

Heynen, N. C., & Lindsey, G. (2003). Correlates of urban forest canopy cover: implications for local public works. Public Works Management & Policy, 8(1), 33-47.

Hodgson, J. A., Kunin, W. E., Thomas, C. D., Benton, T. G., & Gabriel, D. (2010). Comparing organic farming and land sparing: optimizing yield and butterfly populations at a landscape scale. Ecology letters, 13(11), 1358-1367.

IPBES (2019): Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. E. S. Brondizio, J. Settele, S. Díaz, and H. T. Ngo (editors). IPBES secretariat, Bonn, Germany. XXX pages.

IUCN. (2018). A guideline for pollinator-friendly cities. Retrieved from https://www.iucn.org.

Jacke, D., & Toensmeier, E. (2005). Edible forest gardens, volume II: ecological design and practice for temperate-climate permaculture (Vol. 2). Chelsea Green Publishing.

Janssens, I. A., Schauvliege, M., Samson, R., Lust, N., & Ceulemans, R. (1999). De koolstofbalans van drie Vlaamse bosbestanden en voorstellen voor een gericht.

Jennings, S. B., Brown, N. D., & Sheil, D. (1999). Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. Forestry: An International Journal of Forest Research, 72(1), 59-74.

Johnson, N. C., & Wedin, D. A. (1997). Soil carbon, nutrients, and mycorrhizae during conversion of dry tropical forest to grassland. Ecological Applications, 7(1), 171-182.

Jose, S. (2009). Agroforestry for ecosystem services and environmental benefits: an overview. Agroforestry systems, 76(1), 1-10.

Jose, S., Gillespie, A. R., & Pallardy, S. G. (2004). Interspecific interactions in temperate agroforestry. Agroforestry Systems, 61(1-3), 237-255.

Kaufmann, R. K., Zhou, L., Myneni, R. B., Tucker, C. J., Slayback, D., Shabanov, N. V., & Pinzon, J. (2003). The effect of vegetation on surface temperature: A statistical analysis of NDVI and climate data. Geophysical research letters, 30(22).

Kauppi, P. E., Rautiainen, A., Korhonen, K. T., Lehtonen, A., Liski, J., Nöjd, P., ... & Virtanen, T. (2010). Changing stock of biomass carbon in a boreal forest over 93 years. Forest ecology and management, 259(7), 1239-1244.

Kehlenbeck, K., H.S. Arifin, and B.L. Maass, Plant diversity in homegardens in a socio-economic and agro-ecological context, in Stability of Tropical Rainforest Margins. 2007, Springer: The Netherlands. p. 295-317.

Kendall, H. W., & Pimentel, D. (1994). Constraints on the expansion of the global food supply. Ambio, 198-205.

KNMI. (n.d.). KNMI - Klimaat van Nederland. Retrieved April 11, 2020, from https://www.knmi.nl/klimaat.

Konarska, J., Uddling, J., Holmer, B., Lutz, M., Lindberg, F., Pleijel, H., & Thorsson, S. (2016). Transpiration of urban trees and its cooling effect in a high latitude city. International journal of biometeorology, 60(1), 159-172.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. Journal of the American statistical Association, 47(260), 583-621.

Le *et al*, 2008. Sebastien Le, Julie Josse, Francois Husson (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software, 25(1), 1-18. 10.18637/jss.v025.i01.

Lê S, Josse J, Husson F (2008). "FactoMineR: A Package for Multivariate Analysis." Journal of Statistical Software, 25(1), 1–18. doi: 10.18637/jss.v025.i01.

Lee, H. C., McCarl, B. A., & Gillig, D. (2005). The dynamic competitiveness of US agricultural and forest carbon sequestration. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 53(4), 343-357.

Lehmann, L. M., Lysák, M., Schafer, L., & Henriksen, C. B. (2019). Quantification of the understorey contribution to carbon storage in a peri-urban temperate food forest. Urban Forestry and Urban Greening, 45, 126359. https://doi.org/10.1016/j.ufug.2019.06.002.

Levene, H. (1960). Robust testes for equality of variances. In Contributions to Probability and Statistics (I. Olkin, ed.) 278 292. Stanford Univ. Press, Palo Alto, CA. MRO 120709.

Levy, P. E., Hale, S. E., & Nicoll, B. C. (2004). Biomass expansion factors and root: shoot ratios for coniferous tree species in Great Britain. Forestry, 77(5), 421-430.

Lin, B. S., & Lin, Y. J. (2010). Cooling effect of shade trees with different characteristics in a subtropical urban park. HortScience, 45(1), 83-86.

Lugato, E., Panagos, P., Bampa, F., Jones, A., & Montanarella, L. (2014). A new baseline of organic carbon stock in European agricultural soils using a modelling approach. Global change biology, 20(1), 313-326.

Malhi 1999 Malhi, Y. A., Baldocchi, D. D., & Jarvis, P. G. (1999). The carbon balance of tropical, temperate and boreal forests. Plant, Cell & Environment, 22(6), 715-740.

Malhi, Y. A., Baldocchi, D. D., & Jarvis, P. G. (1999). The carbon balance of tropical, temperate and boreal forests. Plant, Cell & Environment, 22(6), 715-740.

Maltamo, M., Bollandsås, O. M., Næsset, E., Gobakken, T., & Packalén, P. (2011). Different plot selection strategies for field training data in ALS-assisted forest inventory. Forestry, 84(1), 23-31.

McBride, S. G., Choudoir, M., Fierer, N., & Strickland, M. S. (2020). Volatile organic compounds from leaf litter decomposition alter soil microbial communities and carbon dynamics. *Ecology*, *101*(10), e03130.

McLauchlan, K. K., Hobbie, S. E., & Post, W. M. (2006). Conversion from agriculture to grassland builds soil organic matter on decadal timescales. Ecological applications, 16(1), 143-153.

Mensah, S., Veldtman, R., & Seifert, T. (2017). Allometric models for height and aboveground biomass of dominant tree species in South African Mistbelt forests. Southern Forests: a Journal of Forest Science, 79(1), 19-30.

Microsoft Corporation. (2018). Microsoft Excel. Retrieved from https://office.microsoft.com/excel.

Mildrexler, D. J., Zhao, M., & Running, S. W. (2011). A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. Journal of Geophysical Research: Biogeosciences, 116(G3).

Mildrexler, D. J., Zhao, M., & Running, S. W. (2011). A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. Journal of Geophysical Research: Biogeosciences, 116(G3).

Morin, X., Fahse, L., Scherer-Lorenzen, M., & Bugmann, H. (2011). Tree species richness promotes productivity in temperate forests through strong complementarity between species. Ecology letters, 14(12), 1211-1219.

Nabuurs, G. J., & Mohren, G. M. J. (1993). Carbon in Dutch forest ecosystems. NJAS wageningen journal of life sciences, 41(4), 309-326.

Nambiar, E. S., & Sands, R. (1993). Competition for water and nutrients in forests. Canadian Journal of Forest Research, 23(10), 1955-1968.

NCA. (2018). Fourth National Climate Assessment: Summary Findings. Retrieved October 2020, from https://nca2018.globalchange.gov/.

NCEP. (n.d.). Climate Prediction Center. Retrieved April 11, 2020, from https://www.cpc.ncep.noaa.go.

Needelman, B. A., Emmer, I. M., Emmett-Mattox, S., Crooks, S., Megonigal, J. P., Myers, D., ... & McGlathery, K. (2018). The science and policy of the verified carbon standard methodology for tidal wetland and seagrass restoration. Estuaries and Coasts, 41(8), 2159-2171.

Nero, B. F., Kwapong, N. A., Jatta, R., & Fatunbi, O. (2018). Tree species diversity and socioeconomic perspectives of the urban (food) forest of Accra, Ghana. Sustainability, 10(10), 3417.

Nytofte, J. L. S., & Henriksen, C. B. (2019). Sustainable food production in a temperate climate–a case study analysis of the nutritional yield in a peri-urban food forest. Urban Forestry & Urban Greening, 45, 126326.

Ong C.K., Corlett J.E., Singh R.P. and Black C.R. 1991. Above and belowground interactions in agroforestry systems. Forest Ecol Manag 45: 45–57.

Opiniepanel (2019). Eenvandaag Opiniepanel Rapport Klimaat. Retrieved from https://eenvandaag.avrotros.nl/fileadmin /user_upload/PDF/Opiniepanel_rapport_Klimaat.pdf.

Palandrani, C., Battipaglia, G., & Alberti, G. (2020). Influence of tree species richness on tree growth and intrinsic water-use efficiency after drought in tree plantations in north-eastern Italy. EUROPEAN JOURNAL OF FOREST RESEARCH.

Parent, S., & Messier, C. (1996). A simple and efficient method to estimate microsite light availability under a forest canopy. Canadian Journal of Forest Research, 26(1), 151-154.

Park, H., Turner, N., & Higgs, E. (2018). Exploring the potential of food forestry to assist in ecological restoration in North America and beyond. Restoration Ecology, 26(2), 284-293.

Patel, N. (2018). Canopy Capture [Mobile application software]. Retrieved from http://play.google.com.

Phalan, B., Balmford, A., Green, R. E., & Scharlemann, J. P. (2011). Minimising the harm to biodiversity of producing more food globally. Food Policy, 36, S62-S71.

Philip, S. Y., Kew, S. F., van der Wiel, K., Wanders, N., & van Oldenborgh, G. J. (2020). Regional differentiation in climate change induced drought trends in the Netherlands. Environmental Research Letters, 15(9), 094081.

Planbureau voor de Leefomgeving. (n.d.). Biodiversiteit en oorzaken van verlies in Europa - Balans van de Leefomgeving 2014 - PBL Planbureau voor de Leefomgeving. Retrieved October 26, 2020, from https://themasites.pbl.nl/balansvandeleefomgeving/jaargang-2014/natuur/biodiversiteit-en-oorzaken-van-verlies-in-europa#:%7E:text=In%20Nederland%20daalde%20de%20biodiversiteit,ongeveer%2015%20procent%20 in%202010.&text=De%20MSA%20geeft%20dus%20weer,milieudruk%20en%20versnippering%20van%20ecosystemen.

Predick, K. I., Gergel, S. E., & Turner, M. G. (2009). Effect of flood regime on tree growth in the floodplain and surrounding uplands of the Wisconsin River. River research and applications, 25(3), 283-296.

Pregitzer, K. S., & Euskirchen, E. S. (2004). Carbon cycling and storage in world forests: Biome patterns related to forest age. Global Change Biology. John Wiley & Sons, Ltd. https://doi.org/10.1111/j.1365-2486.2004.00866.x

QGIS.org (2020). QGIS Geographic Information System. Open Source Geospatial Foundation Project. https://QGIS.org.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

R. B. Brown (September 2007). "Soil Texture" (PDF). Agronomy Fact Sheet Series: Fact Sheet SL-29. Cornell University, Department of Crop and Soil Sciences. Retrieved May 2, 2017.

Remy, E., Wuyts, K., Boeckx, P., Ginzburg, S., Gundersen, P., Demey, A., ... & Verheyen, K. (2016). Strong gradients in nitrogen and carbon stocks at temperate forest edges. Forest Ecology and Management, 376, 45-58.

Riolo, F. (2019). The social and environmental value of public urban food forests: The case study of the Picasso Food Forest in Parma, Italy. Urban Forestry & Urban Greening, 45, 126225.

Ruotsalainen, R., Pukkala, T., Kangas, A., Vauhkonen, J., Tuominen, S., & Packalen, P. (2019). The effects of sample plot selection strategy and the number of sample plots on inoptimality losses in forest management planning based on airborne laser scanning data. Canadian Journal of Forest Research, 49(9), 1135-1146.

Schafer, L. J., Lysák, M., & Henriksen, C. B. (2019). Tree layer carbon stock quantification in a temperate food forest: A peri-urban polyculture case study. Urban Forestry and Urban Greening, 45, 126466. https://doi.org/10.1016/j.ufug.2019.126466.

Schelhaas, M., & Clerkx, A. P. P. M. (2015). Het Nederlandse bos in cijfers: resultaten van de 6e Nederlandse Bosinventarisatie. Vakblad Natuur Bos Landschap, 12(111), 23-27.

Schulte, R. P. O., Diamond, J., Finkele, K., Holden, N. M., & Brereton, A. J. (2005). Predicting the soil moisture conditions of Irish grasslands. Irish Journal of Agricultural and Food Research, 95-110.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591.

Sharma, S. K., Telfer, M., Phua, S. T., & Chandler, H. (2012). A pragmatic method for estimating greenhouse gas emissions from leakage for Improved Forest Management projects under the Verified Carbon Standard. Greenhouse Gas Measurement and Management, 2(1), 22-32.

Shen, X., Liu, B., Zhou, D., & Lu, X. (2016). Effect of grassland vegetation on diurnal temperature range in China's temperate grassland region. Ecological Engineering, 97, 292-296.

Sigurdsson, B. D., Magnusson, B., Elmarsdottir, A., & Bjarnadottir, B. (2005). Biomass and composition of understory vegetation and the forest floor carbon stock across Siberian larch and mountain birch chronosequences in Iceland. Annals of Forest Science, 62(8), 881-888.

Sikkema & Nabuurs Sikkema, R., & Nabuurs, G. J. (1994). Forests and wood in the carbon cycle: Study on the CO [sub 2] removal options of four types of forests: oak/beech, spruce, long-cycle poplar (15 year) and short-cycle poplar (5 year).

Simonin, K. A., Link, P., Rempe, D., Miller, S., Oshun, J., Bode, C., ... & Dawson, T. E. (2014). Vegetation induced changes in the stable isotope composition of near surface humidity. Ecohydrology, 7(3), 936-949.

Stahl, D. A., & Urbance, J. W. (1990). The division between fast-and slow-growing species corresponds to natural relationships among the mycobacteria. Journal of bacteriology, 172(1), 116-124.

Stephenson, N. L., Das, A. J., Condit, R., Russo, S. E., Baker, P. J., Beckman, N. G., ... & Alvarez, E. (2014). Rate of tree carbon accumulation increases continuously with tree size. Nature, 507(7490), 90-93.

Sullivan, P. F., Ellison, S. B., McNown, R. W., Brownlee, A. H., & Sveinbjörnsson, B. (2015). Evidence of soil nutrient availability as the proximate constraint on growth of treeline trees in northwest Alaska. Ecology, 96(3), 716-727.

Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot.

Teuling, A.J., Seneviratne, S.I., Stöckli, R. *et al.* (2010) Contrasting response of European forest and grassland energy exchange to heatwaves. Nature Geoscience. 3: 722- 727.

Trumbore, S. E. (1993). Comparison of carbon dynamics in tropical and temperate soils using radiocarbon measurements. Global Biogeochemical Cycles, 7(2), 275-290.

Tscharntke, T., Clough, Y., Wanger, T. C., Jackson, L., Motzke, I., Perfecto, I., ... & Whitbread, A. (2012). Global food security, biodiversity conservation and the future of agricultural intensification. Biological conservation, 151(1), 53-59.

Ullah, M. R., & Al-Amin, M. (2012). Above-and below-ground carbon stock estimation in a natural forest of Bangladesh. Journal of forest science, 58(8), 372-379.

UN. (2018). 2018 United Nations List of Protected Areas. Retrieved from https://www.sperp.org.

UNFCCC. (2010, November). A/R Methodological Tool: Calculation of the number of sample plots for measurements within A/R CDM project activities (Version 2.1.0). Retrieved from https://cdm.unfccc.int/methodologies/ARmethodologies/tools/ar-am-tool-03-v2.1.0.pdf/history_view.

UNFCCC. (2013). Methodological Tool: Estimation of carbon stocks and change in carbon stocks of trees and shrubs in A/R CDM project activities. (04.2). Retrieved from https://cdm.unfccc.int/methodologies/ARmethodologies/tools/ar-am-tool-14-v4.2.pdf.

Urban, D. L., & Shugart, H. H. (1992). Individual-based models of forest succession. Plant succession: theory and prediction, 249-292.

Valverde, T., & Silvertown, J. (1997). Canopy closure rate and forest structure. Ecology, 78(5), 1555-1562.

Van der Meijden, R. (2005). Heukels' Flora van Nederland (23e druk. ed.). Groningen/Houten, The Netherlands: Noordhoff.

van Noordwijk, M., Bayala, J., Hairiah, K., Lusiana, B., Muthuri, C., Khasanah, N. M., & Mulia, R. (2014). Agroforestry solutions for buffering climate variability and adapting to change. Climate change impact and adaptation in agricultural systems. CAB-International, Wallingford, 216-232.

Veenman, S., Liefferink, D., & Arts, B. (2009). A short history of Dutch forest policy: The 'de-institutionalisation'of a policy arrangement. Forest Policy and Economics, 11(3), 202-208.

Verra. (n.d.). Verified Carbon Standard. Retrieved November 5, 2020, from https://verra.org/methodologies/.

Verschuyl, J., Clark, L., & Loehle, C. (2018). Predicting shrub biomass and current annual growth from field measurements in the Oregon Coast Range. Northwest Science, 92(1), 9-17.

Vieilledent, G., Vaudry, R., Andriamanohisoa, S. F., Rakotonarivo, O. S., Randrianasolo, H. Z., Razafindrabe, H. N., ... & Rasamoelina, M. (2012). A universal approach to estimate biomass and carbon stock in tropical forests using generic allometric models. Ecological Applications, 22(2), 572-583.

Voedselboskaart, 2020. URL: http://www.voedselboskaart.nl.

Von Avenarius, A., Devaraja, T. S., & Kiesel, R. (2018). An Empirical Comparison of Carbon Credit Projects under the Clean Development Mechanism and Verified Carbon Standard. Climate, 6(2), 49.

Wang, C., Zhao, C., Xu, Z., Wang, Y., & Peng, H. (2013). Effect of vegetation on soil water retention and storage in a semi-arid alpine forest catchment. Journal of arid land, 5(2), 207-219.

Warren, D. R., Keeton, W. S., Bechtold, H. A., & Rosi-Marshall, E. J. (2013). Comparing streambed light availability and canopy cover in streams with old-growth versus early-mature riparian forests in western Oregon. Aquatic sciences, 75(4), 547-558.

Watt, A. D., Stork, N. E., Eggleton, P., Srivastava, D., Bolton, B., Larsen, T. B., ... & Bignell, D. E. (1997). Impact of forest loss and regeneration on insect abundance and diversity. Forests and insects, 273-286.

Wheater, H., & Evans, E. (2009). Land use, water management and future flood risk. Land use policy, 26, S251-S264.

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2'. Create Elegant Data Visualisations Using the Grammar of Graphics. Version, 2(1), 1-189.

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics 1:80–83.

Wilson, M. V. (2015). Simple random sampling in the field. Retrieved March 25, 2020, from https://oregonstate.edu/instruct/bot440/wilsomar/Content/SRS.htm.

WWF. (2014). Living Planet Report 2014. Retrieved from https://www.wwf.or.jp/activities/data/WWF_LPR_2014.pdf.

Zenner, E. K., & Hibbs, D. E. (2000). A new method for modeling the heterogeneity of forest structure. Forest ecology and management, 129(1-3), 75-87.

Appendix 1: Information of 21 included food forests

Table A1: Information of 21 food forests included in this study, including 42 zones in total. Not all information could be shown in one table, other information can be found in *Table A2*. **Food forest** = name of the food forest. **Zone** = Coding of the zone, the first letter referred to the food forest, the second letter referred to the zone (A, B, C or D). **Detailed FLU** = Zone former land use. **FLU =** Former land use category in which the zone has been assigned: *forest, arable land* or *grassland*. **Planting** = The year of the first planting event. All planting events in the period November – March has been assigned to the year of the growing season. **Age** = The age of the zone in growing seasons. * = Food forest Groengenoten was divided in two zones based on its former land use, but both FLU's were classified in category *grassland*.

Food Forest	Zone	Detailed FLU	FLU	Planting	Age
De Overtuin	**AA**	Arboretum	Forest	2019	2
De Overtuin	**AB**	Arboretum	Forest	2019	2
De Overtuin	**AC**	Arboretum	Forest	2019	2
De Overtuin	**AD**	Arboretum	Forest	2019	2
Houtrak	**BA**	Arable Land	Arable Land	2017	4
MijnStadstuin	**CA**	Arable Land	Arable Land	2016	5
MijnStadstuin	**CB**	Arable Land	Arable Land	2016	5
Thuishaven	**DA**	Arable Land	Arable Land	2018	3
Droevendaal	**EA**	Orchard	Forest	2019	2
Voedselrijk	**FA**	Production Forest	Forest	2019	2
Voedselrijk	**FB**	Production Forest	Forest	2019	2
Voedselrijk	**FC**	Production Forest	Forest	2019	2
Voedselrijk	**FD**	Production Forest	Forest	2019	2
Eemvallei_Zuid	**GA**	Arable Land	Arable Land	2020	1
Eemvallei_Zuid	**GB**	Arable Land	Arable Land	2020	1
Den Food Bosch	**HA**	Maize Pasture	Arable Land	2017	4
Schijndel Boschweg	**IA**	Arable Land	Arable Land	2019	2
Hardekamp	**JA**	Ryegrass field	Grassland	2019	2
Schijndel Hardekamp	**JB**	Ryegrass field	Grassland	2020	1
Schijndel Hardekamp	**JC**	Ryegrass field	Grassland	2021	0
Ketelbroek	**KA**	Maize Pasture	Arable Land	2015	6
Ketelbroek	**KB**	Maize Pasture	Arable Land	2009	12
Groengenoten	**LA**	Horse Field*	Grassland	2019	2
Groengenoten	**LB**	Horse Field*	Grassland	2019	2
Sualmana	**MA**	Sheep Field	Grassland	1999	22
Vlaardingen	**NA**	Recreation Field	Grassland	2015	6
Vlaardingen	**NB**	Forest	Forest	2015	6
Vlaardingen	**NC**	Recreation Field	Grassland	2015	6
Benthuizen	**OA**	Arable Land	Arable Land	2020	1
Benthuizen	**OB**	Arable Land	Arable Land	2018	3
De Stomp	**PA**	Production Forest	Forest	2019	2
Kreilerwoud	**QA**	Horse Field	Grassland	2017	4
Roggebotstaete	**RA**	Tree Nursery	Forest	2016	5
Roggebotstaete	**RB**	Tree Nursery	Forest	2016	5
D'Ekkers	**SA**	Ryegrass field	Grassland	2020	1
D'Ekkers	**SB**	Ryegrass field	Grassland	2020	1

Breedenbroek	**TA**	Barly Field	Arable Land	2020	1
Breedenbroek	**TB**	Barly Field	Arable Land	2021	0
Lekker Landgoed	**UA**	Herb-Rich Grassland	Grassland	2016	5
Lekker Landgoed	**UB**	Arable Land	Arable Land	2017	4
Lekker Landgoed	**UC**	Herb-Rich Grassland	Grassland	2018	3
Lekker Landgoed	**UD**	Herb-Rich Grassland	Grassland	2019	2
Reference Groengenoten	**VB**		Forest		
Reference Ketelbroek	**WB**		Forest		
Reference Thusihaven	**XB**		Forest		

Table A2: Information of 21 food forests included in this study, including 42 zones in total. Not all information could be shown in one table, other information can be found in *Table A1*. **Zone** = Coding of the zone, the first letter referred to the food forest, the second letter referred to the zone (A, B, C or D). **Structural Composition** = Structural composition of the zone, either *dispersed* or *Alleys*. **Altitude** = Category of zone altitude, either *normal* or *raised* (mounds). **Soil Texture** = Soil texture class, based on the soil texture triangle: Loam, Sand or Clay. **Carbon** = Aboveground carbon stock of the zone in Mg Co_2 ha^{-1}. **Canopy =** Average percentage of the canopy closure. **Temp** = Relative difference in temperature between inside and outside measurements in % of outside temperature. **Hum** = Relative difference in temperature between inside and outside measurements in % of outside humidity. *N.A.* = data not available.

Zone	Structural Composition	Altitude	Soil Texture	Carbon	Canopy	Temp	Hum
AA	Dispersed	Normal	Loam	43,57	76,43	-28,04	-9,83
AB	Alleys	Normal	Loam	2,03	56,00	-27,33	16,67
AC	Alleys	Normal	Loam	18,46	58,10	-49,11	44,56
AD	Dispersed	Normal	Loam	32,05	63,77	-49,42	58,20
BA	Alleys	Normal	Clay	0,13	0,00	*N.A.*	*N.A.*
CA	Dispersed	Raised	Loam	0,49	*N.A.*	*N.A.*	*N.A.*
CB	Dispersed	Normal	Loam	0,46	*N.A.*	*N.A.*	*N.A.*
DA	Dispersed	Normal	Clay	0,06	0,00	-35,85	64,09
EA	Alleys	Normal	Sand	2,29	10,77	-34,25	79,47
FA	Dispersed	Normal	Sand	159,31	58,20	*N.A.*	*N.A.*
FB	Dispersed	Normal	Sand	90,46	58,13	*N.A.*	*N.A.*
FC	Dispersed	Normal	Sand	194,67	66,10	*N.A.*	*N.A.*
FD	Dispersed	Normal	Sand	95,19	71,00	*N.A.*	*N.A.*
GA	Alleys	Normal	Clay	0,00	0,00	-20,60	23,09
GB	Dispersed	Normal	Clay	0,00	0,00	-19,84	9,73
HA	Dispersed	Normal	Sand	0,00	3,10	-9,28	0,21
IA	Dispersed	Normal	Sand	0,00	0,00	-18,35	68,11
JA	Dispersed	Normal	Loam	0,04	0,00	-27,22	35,83
JB	Alleys	Normal	Loam	0,02	0,00	-26,43	28,66
JC	Alleys	Normal	Loam	0,00	0,00	3,56	-43,49
KA	Alleys	Normal	Loam	4,58	35,10	-55,00	216,09
KB	Dispersed	Normal	Loam	15,87	59,00	-18,75	32,31
LA	Alleys	Normal	Sand	0,01	6,33	-36,87	164,99
LB	Alleys	Normal	Sand	0,01	7,10	-29,32	131,47
MA	Dispersed	Normal	Sand	37,16	82,47	*N.A.*	*N.A.*
NA	Dispersed	Raised	Loam	3,93	10,23	-10,24	-4,16

NB	Dispersed	Normal	Loam	113,20	81,23	-32,20	25,47
NC	Dispersed	Normal	Loam	4,53	36,77	*N.A.*	*N.A.*
OA	Alleys	Normal	Clay	0,09	0,57	-27,90	67,82
OB	Dispersed	Normal	Clay	0,30	0,00	-38,92	124,14
PA	Alleys	Normal	Loam	169,60	86,35	-19,46	-3,04
QA	Alleys	Raised	Loam	0,64	0,00	-25,48	27,87
RA	Alleys	Normal	Sand	9,58	48,37	-16,75	-1,93
RB	Alleys	Normal	Loam	7,19	71,23	-31,50	9,48
SA	Dispersed	Normal	Sand	0,01	0,00	*N.A.*	*N.A.*
SB	Alleys	Normal	Sand	0,00	0,00	*N.A.*	*N.A.*
TA	Dispersed	Normal	Sand	0,04	0,00	-24,80	49,63
TB	Alleys	Normal	Sand	0,00	0,00	-18,26	32,16
UA	Dispersed	Raised	Clay	1,25	19,53	-11,18	1,29
UB	Dispersed	Normal	Clay	0,36	0,00	-26,52	19,24
UC	Dispersed	Normal	Clay	0,12	0,00	-7,37	-6,23
UD	Dispersed	Normal	Clay	0,01	0,00	-18,48	0,45
VB	Dispersed	Normal	Sand	111,01	62,30	-19,55	61,19
WB	Dispersed	Normal	Loam	263,60	93,70	-49,14	151,35
XB	Dispersed	Normal	Clay	217,95	90,70	-59,79	147,51

Appendix 2: Protocols of the fieldwork

In this section, the protocols that have been used during the fieldwork are given, including:

- Protocol on sampling plot selection in food forests.
- Protocol on sampling plot selection in reference forests.
- Protocol on measurements within the 10x10m sampling plot.
- Protocol on temperature and humidity measurements.
- Soil Texture Triangle used to determine soil texture classes (*Figure A1*).
- List of tree and shrub species commonly planted in food forests used to determine whether plants are shrubs or trees (*Table A3*).

Appendix 2.1. Protocol for random selection of sampling plots in food forests

The following protocol is to determine the location of sampling plots in the food forests. This protocol has also been used in other studies part of the NMPF.

1. Find the concerning cadastral parcel in QGIS (using the PDOK database).
2. Geo-reference the food forest design / map including landscape elements with the cadastral parcel.
3. Use 'Create grid (in Vector > research tools) to place a grid over the food forest parcel. Use the layer extent of the food forest as grid extent and use rectangle 10x10m as grid type.
4. Determine whether the food forest must be divided in zones (in consultation with the NMPF-coordinator and/or food forest manager).
5. Indicate the zones by splitting drawing the zone boundaries and splitting up the grid.
6. Remove all grids that lay partly outside of the food forests manually. Remove grids that lay in two zones manually as well.
7. Determine the number of sampling plots to select:
    a. If the food forest is not divided in zones, the number of sampling plots depends on the food forest size. Multiply the total area in hectares with 3 to obtain the number of sampling plots. A minimum 3 and a maximum of 6 plots will be handled. *Example:* 1.75*3=5.25 (rounded to 5) sampling plots will be selected in a food forest with a size of 1.75ha.
    b. If the food forest is divided in zones, the fixed number of 3 sampling plots per zone will be used. *Example:* 3x3=9 sampling plots will be selected in a food forest with 3 zones. There is one exception: If a zone consists of 5 grids or less, only 2 sampling plots will be selected.
    c. Blue measuring points (that are used for NMPF soil measurements before) will automatically be selected as sampling point for this research as well. The number of blue points in a food forest or zone must be deducted from the points to be selected). *Example:* 3 sampling plots must be allocated to a certain zone and 1 blue point is located in this zone. The blue point will automatically be selected. The other 2 sampling points will be selected randomly (see Step 8).
8. Use the tool 'Random selection' (in Vector > research tools > random selection points) to select the sampling points randomly. If the food forest is undivided, one selection will be performed for the entire grid. If the food forest is divided, one selection per zone will be performed.
9. If the selected grid is 'impossible or undesirable', one grid westwards is selected alternatively. If this grid is 'impossible or undesirable' as well, one grid northwards (relatively to the original grid) is selected. Or, as last option, one grid eastwards or

southwards, in this order. A grid is classified as 'impossible or undesirable' when a pond, paved path or building is located in the grid.

10. If all sampling plots are selected, save the random selection as new layer (export > save selected features as…).
11. Expand the attribute table with a new field for the labels (text string, length = 5), with title ID_NMVB. Use the labels in the map. Make sure that each zone has a specific color, to clearly visualise to which zone a certain sampling plot belongs.
12. Arrange all sampling plots from left to right and from top to bottom in the attribute table, by clicking on *column left* and *column top*. Label all sampling plots in this order with a capital letter to indicate the food forest, a lowercase letter to indicate the zone and a number to indicate the sampling plot. *Example:* Sa1, Sa2, Sa3, Sb1, Sb2, Sb3, etc.
13. Export all layers and the parcel boundaries as kml-file by using the MMQGIS plugin. Make sure that all plots are labeled with the ID_NMVB field.
14. Open the kml-file in Google Earth (or any other software) and export it to your mobile device.
15. If a selected plot turns out to be 'impossible or undesirable' in the field because of water, paved paths or buildings, the plot could be moved westwards, northwards, eastwards, or southwards in this order.

Appendix 2.2. Protocol for random selection of sampling plots in reference forests

In the reference forests, only one sampling plot is selected. The following, less-time consuming protocol has been used to locate this sampling plot:

1. Start at the southeastern corner of the forest parcel (if this is not possible, determine another starting point which could be found back easily).
2. Set a timer at 1 minute and start walking northwestwards (if started at an alternative starting point, walk 1 minute towards the center of the parcel (use google maps to determine walking direction).
3. Stop walking after 1 minute and start a stopwatch.
4. Stop the stopwatch blindly at a random moment and notate the first two decimals shown by the stopwatch.
5. Walk the number of meters shown by the first decimal northwards and the number of meters shown by the second decimal westwards. This point is the northeastern corner of the sampling plot. (For example: If the stopwatch has been stopped at 00:03:56, the northeastern corner of the sampling plot is located 5 meters northwards and 6 meters westwards).
6. Determine the other corners of the sampling plot by walking ten meters southwards (southeastern corner) and westwards (northwestern corner) and ten meters southwards from the northwestern corner.
7. If the determined sampling plot could not be used (because of a paved road, pond or building), the end point of the random selection methods (end of step 5) is used as northwestern corner of the sampling plot instead (*i.e.* the entire sampling plot is moved 10 meters eastwards). When this is not possible as well, the end point of step 5 is used as southwestern corner instead. When this is not possible as well, the end point of step 5 is used as southeastern corner instead. If all options are not possible, this protocol must be performed again from the beginning.

Appendix 2.3. Protocol for all measurements within the 10x10 meter sampling plot

Supplies: Bamboo sticks (4 per sampling plots), measuring tape of 10 metres, compass, inclinometer, application PlantNet, list of food forest shrub and tree species (NMPF), Heukels' Flora van Nederland (or other identification books), thermohygrometer.

The following protocol must be considered to determine the sampling plots:
1. The location of the center of the sampling plot was found through its GPS-location.
2. From the center of the sampling plot, one walked five meters northward and five meters eastward to determine the northeastern corner of the plot. A steel bar was placed at this spot and labeled with the code of the plot.
3. From the northeastern corner, the other corners were determined by walking ten meters southwards and westwards and walking ten meters southwards from the northwestern corner. A metal herring was inserted in the soil at these three corners.

The following measurements must be taken for all individual trees and shrubs within the 10x10m sampling plot. Only those trees of which *the stem* was entirely located within the plot were included. Only shrubs of which the *center of the basal area* was entirely located within the plot were included.
1. Determine the species of the individual using Heukels' Flora or other reference works (Plantnet could not be used for species determination).
2. Determine the correctness of the application Plantnet, by allocation one of the following four categories:
    a. The actual species is at the top of the list provided by Plantnet, having a score of at least 3.00.
    b. The actual species is at the top of the list provided by Plantnet, having a score of less than 3.00.
    c. The actual species is at the second or third place of the list provided by Plantnet.
    d. The actual species is not present in the top three of the list provided by Plantnet.
3. Determine whether the individual is planted or spread wildly.
4. Determine whether the individual is a tree or a shrub. The list of shrub and tree species provided by the NMPF will be used, even when the growth form of the concerning individual suggests the other category. If the species is not included in the list, the species will be allocated to the category of related species. In case of doubt, the coordinator of the NMPF must be consulted.

The following measurements must be taken for all trees within the sampling plot:
1. Determine whether the diameter of the stem is at least 5 mm. If the diameter is not, determine the number of individuals of this species in the entire sampling plot. If the diameter is at least 5 mm, continue the measurements following the next steps.
2. Determine the maximum height of the tree in centimeters, using an inclinometer. The height of trees less than 2 meters is determined using measuring tape instead of an inclinometer.
3. Determine the diameter of the stem at 130 centimeters aboveground. If the stem is branched off at this height, determine stem diameter at 60 centimeters aboveground instead. If the steam is branched of at 60 centimeters as well, determine stem diameter just below the bottom branch (notate measuring height as well in the last scenario).

The following measurements must be taken for all shrubs within the sampling plot:
1. Determine the maximum height of the shrub in centimeters, using an inclinometer. The height of shrubs less than 2 meters is determined using measuring tape instead of an inclinometer.

2. Determine the diameter of the crown in centimeters in two perpendicular directions. Only the crown area located within the sampling plot will be included.
3. Determine the number of stems with a diameter of more than 5mm.
4. Determine the circumference of the three thickest stems of the shrub at 30 centimeters aboveground.

Appendix 2.4. Protocol for temperature and humidity measurements

1. An automatic data logger will be used to determine humidity and temperature. The logger must be set to six measurements per hour (one measurement every ten minutes).
2. The outside logger must be placed at least 10 meters outside the food forest at an unpaved surface. Ideally, the outside logger will be placed in the roadside. The logger should not be shadowed by any trees or buildings.
3. The inside logger will be moved along all sampling plots. During the displacement of the logger, the measurements will not be paused. Measurements taken during displacements must be removed manually after the data was extracted from the logger. The displacement events should be as short as possible, and the logger will be carried in an empty bag to prevent it for being heated up. During the logger is placed at the center of the sampling plot in the vegetation, even when this location is shadowed by any tree or shrub.
4. At least 3 measurements are necessary per location, which means that the data logger must be placed at the sampling plot for at least 30 consecutive minutes.

Appendix 2.5 Protocol for zone selection

Since the selection of plots is standardised in the national monitoring program, the decision whether to split up food forests or not should ideally be standardised either. However, zone selection will always be subjective. In this research, the zone selection was limited to the following aspects. If a food forest had intern differences in one or multiple of these aspects, this food forest could be divided into multiple zones (which does not mean that this was always been done):

- Age of the food forest, which is determined as the number of growing seasons since the first major planting event. The growing season starts in March each year.
- Former land use, which is limited to three former land use categories (see *section 3.3*).
- Soil texture, which is limited to the three soil texture classes (see *section 3.3*).
- Elevation. The minimum difference in elevation should be 1m.
- Structural composition.

Especially the last criterion is food for thought, since the vegetation structure in a food forest is commonly very heterogeneous. The main reason to divide a food forest into two zones based on structural composition was the difference between production-oriented, so-called ‘rational food forests’ and more forest-looking ‘romantic food forests’. Although these terms are common for insiders, the terms ‘alleys’ (for rational food forests), and ‘dispersed’ (for romantic food forests) will be used in this research instead. Other differences in vegetation structure could also be a reason for applying zones. This was always discussed with the management of the national monitoring programming to pursue maximum objectivity.

After the criteria for zone selection were drawn up, it was necessary to determine whether these situations were suitable for a zone or not. Even with these restrictions and criteria, it was impossible to capture a subjective decision into an objective protocol, although it might help determining whether zones should be implemented in future or not. The following conditions should be met to split up a food forest into zones:

- The zones are clearly distinguishable. If there is a gradient from one condition to another, it is impossible to determine whether the zone border should be located.
- The zone has a minimum size of 10 complete grids. Since the minimum sample effort was determined at three plots per zone, this minimum size is necessary to select plots randomly.

As in the one-zone food forests, the plot selection was made using QGIS (QGIS, 2020). In collaboration with food forest owners, a map with the zones was drawn up and georeferenced to the parcel in QGIS. In this way, the exact borders of the zones were determined. The plots that were located at the border of two zones were deleted from the grid. The grid of the food forest was split up into a grid for all zones each, after which three measuring plots were randomly selected by the random selection tool in QGIS (see *Appendix 3* for the entire protocol). Since the amount of measuring plots automatically equalled or transcended the maximum number of plots selected in one-zone food forests, the number of plots per zone were not size dependent. Lowering the minimum number of plots per zone would have increased the chance of coincidental observations and was therefore not applied, while enhancing the number of plots would have been made the sampling effort less accessible.

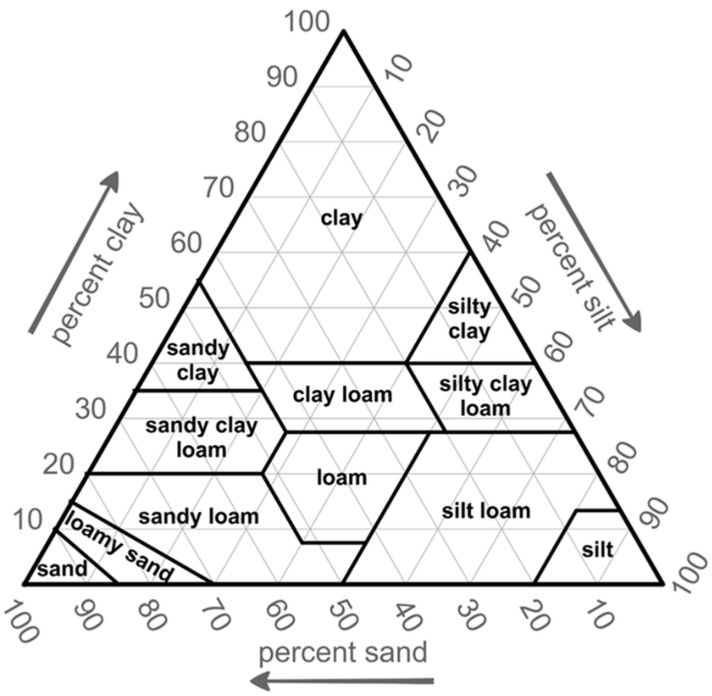


Figure A1: Soil Texture Triangle (Groenendyk *et al.,* 2015). Based on this triangle, food forests were divided into three soil texture classes (sand, loam and clay).

Table A3: List of tree and shrub species commonly planted in food forests. This list has been used to determine whether tree or shrub equations were used.

Tree species	Shrub species
Acer saccharum	*Acca sellowiana*
Acer (other species)	*Amelanchier lamarckii*
Araucaria araucana	*Arbutus unedo*
Asimina triloba	*Berberis vulgaris*
Betula pendula (or *pubscens)*	*Chaenomeles cathayensis*
Carya ilinoinensis	*Chaenomeles japonica*
Carya illinoiensis x laciniosa	*Cornus mas*
Carya laciniosa	*Corylus avellana*
Carya tomentosa	*Cydonia oblonga*
Castanea sativa	*Eleagnus multiflora*
Cercis siliquastrum	*Eleagnus umbellate*
Citrus (multiple species)	*Eleagnus x ebbingei*
Corylus colurna	*Ficus carica*
Crataegus germanica	*Hibiscus syriacus* (or other species)
Crataegus schraderiana	*Hippophae rhamnoides*
Diospyros kaki	*Lonicera caerulea*
Diospyros lotus	*Lycium barbarum*
Diospyros virginiana	*Mahonia aquifolium*
Eribotrya japonica	*Myrica gale*
Ginkgo biloba	*Poncirus trifoliata*
Gleditsia triacanthos	*Ribes* (other species)
Halesia carolina	*Ribes nidigrolaria*
Hovenia dulcis	*Ribes nigrum*
Juglans ailantifolia	*Ribes odoratum*
Juglans ailantifolia x cinerea	*Ribes rubrum*
Juglans cinerea	*Rubus fructicosus*
Juglans nigra	*Rubus idaeus*
Juglans regia	*Rubus ideaus x fruticosus* (tayberry, loganberry, etc.)
Magnolia (multiple species)	*Rubus phoenicolasius*
Malus domestica	*Sambucus nigra*
Morus alba	*Staphylea pinnata*
Morus nigra	*Tilia cordata* (if planted in a hedge)
Pinus koraiensis	*Ugni molinea*
Pinus pinea	*Vaccinium corymbosum*
Prunus armeniaca	*Viburnum lentago*
Prunus avium	*Xanthocera sorbifolium*
Prunus cerasus	*Zanthoxylum simulans* (or *piperitum)*
Prunus domestica	
Prunus dulcis	
Prunus persica	
Prunus persica nucipersica	
Prunus salicina (or *triflora)*	
Pyrus communis	
Pyrus pyrifolia	
Quercus (multiple species)	
Robinia pseudoacacia	
Sorbopyrus auriculata	
Sorbus aucuparia	
Sorbus domestica	
Tilia cordata	
Toona sinensis	
Torreya grandis grandis	

Appendix 3: Assumption Tests

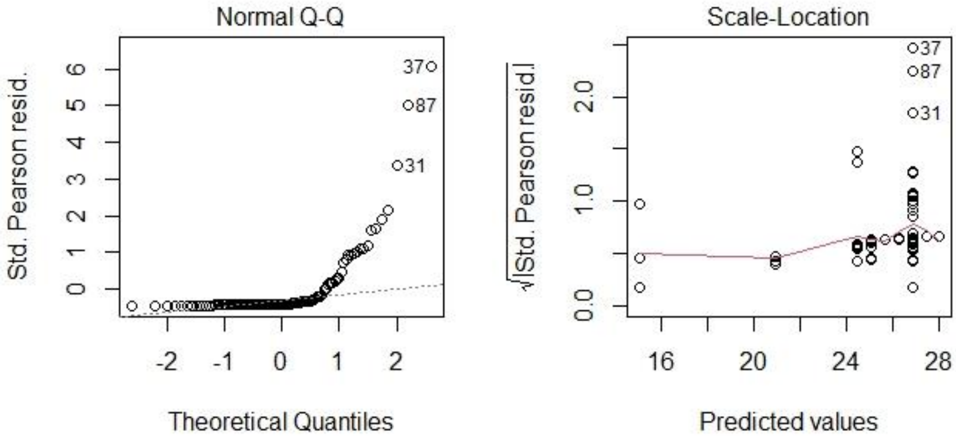


Figure A2: Diagnostic plots belonging to the generalised linear model: glm(Carbon ~ Age). Data was not normal (left graph) and heteroscedastic (right graph). n = 130. *Exported from Rstudio.*

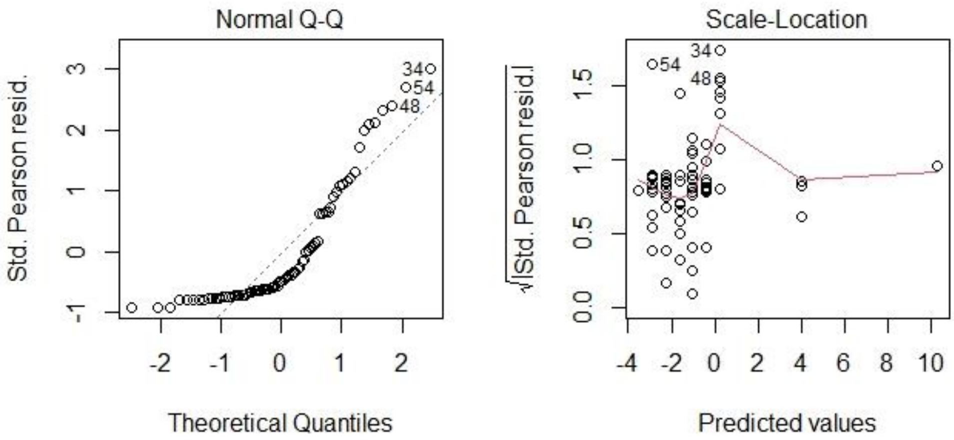


Figure A3: Diagnostic plots belonging to the generalised linear model: glm(Carbon ~ Age), including all food forests on former arable lands and grasslands. Data was not normal (left graph) and heteroscedastic (right graph). n = 95. *Exported from RStudio.*

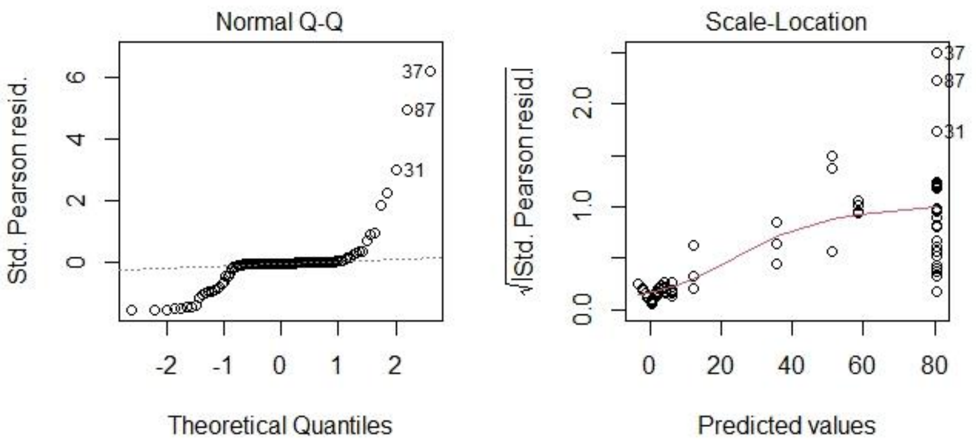


Figure A4: Diagnostic plots belonging to the generalised linear model: glm(Carbon ~ Age * FLU).
Data was not normal (left graph) and heteroscedastic (right graph). n = 130. *Exported from RStudio.*

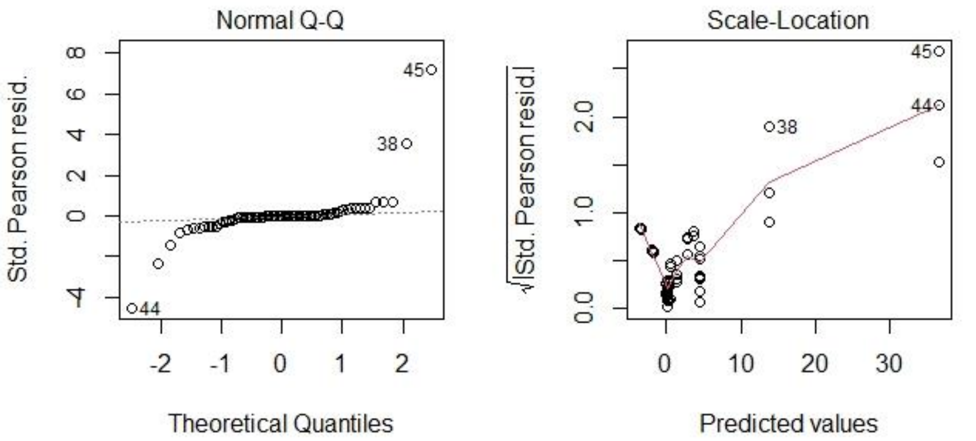


Figure A5: Diagnostic plots belonging to the generalised linear model: glm(Carbon ~ Age * Soil Texture).
Data was not normal (left graph) and heteroscedastic (right graph). n = 130. *Exported from Rstudio.*

Table A4: Results of assumption tests concerning ANOVA's and T-tests. Results of both linearity tests (Shapiro-Wilk) and homoscedastisty tests (Levene) were geven, including the selected significance test and the corresponding test value.

Tested Hypothesis	Tested Data	Shapiro-Wilk Test	Levene's Test	Used Hypothesis Test	Test Value
Difference in carbon stock between FLU types	lm(Carbon ~ FLU)	2.2e-16	4.558e-10		
Difference in carbon stock between FLU types after log transformation	lm(CarbonLog ~ FLU)	3.561e-07	7.241e-06	Kruskal-Wallis significance test	p = 5.574e-13
Difference in carbon stock between soil types	lm(Carbon ~ Soiltype)	2.2e-16	0.01193		
Difference in carbon stock between soil types after log transformation	lm(CarbonLog ~ Soiltype)	1.537e-08	2.176e-05	Kruskal-Wallis significance test	p = 0.0006926
Difference in carbon stock between structural compositions	lm(Carbon ~ Composition)	2.2e-16	0.009534		
Difference in carbon stock between structural compositions after log transformation	lm(CarbonLog ~ Composition)	2.27e-10	0.0043	Mann-Whitney U Test	p = 0.03242
Difference in carbon stock between structural compositions, only for food forests with both compositions	lm(Carbon ~ Composition)	1.435e-10	0.1974		
Difference in carbon stock between structural compositions after log transformation, only for food forests with both compositions	lm(CarbonLog ~ Composition)	4.517e-09	0.2421	Mann-Whitney U Test	p = 0.02092
Difference in temperature inside and outside the food forest		Ins: 0.1447 Out: 0.3697		T-test for two paired samples	Est: -10.14146 p = 1.123e-08
Difference in humidity inside and outside the food forest		Ins: 0.1242 Out: 0.1144		T-test for two paired samples	Est:12.03301 p = 1.144e-05
Difference in carbon stock between food forests and reference forests		REF: 0.5648 FF: < 2.2e-16		Mann-Whitney U Test	P = 0.004813 Est: 216.59
Difference in carbon stock between food forests with FLU *forest* and reference forests		FF FLU Forest: 2.773e-06		Mann-Whitney U Test	p = 0.01253 Est: 135.67

Appendix 4: Supplementary Results

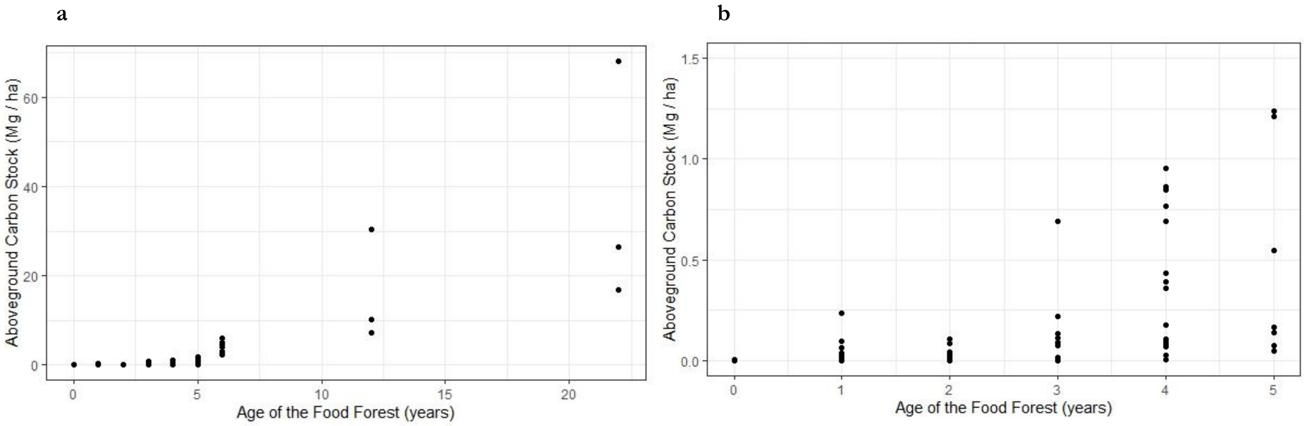


Figure A6: Food forest aboveground cabon stock versus food forest age on former arable lands and grasslands. Scatter points representate sampling plots). Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}, age is expressed in years. ANOVA and Kruskal-Wallis tests were used to describe the regression between both variables **a)** Aboveground carbon stock versus food forest age for food forest with FLU *arable land* and *grassland* (n = 81). **b)** Aboveground carbon stock versus food forest age for food forest with FLU *arable land* and *grassland* (n = 95). *Exported from RStudio,*

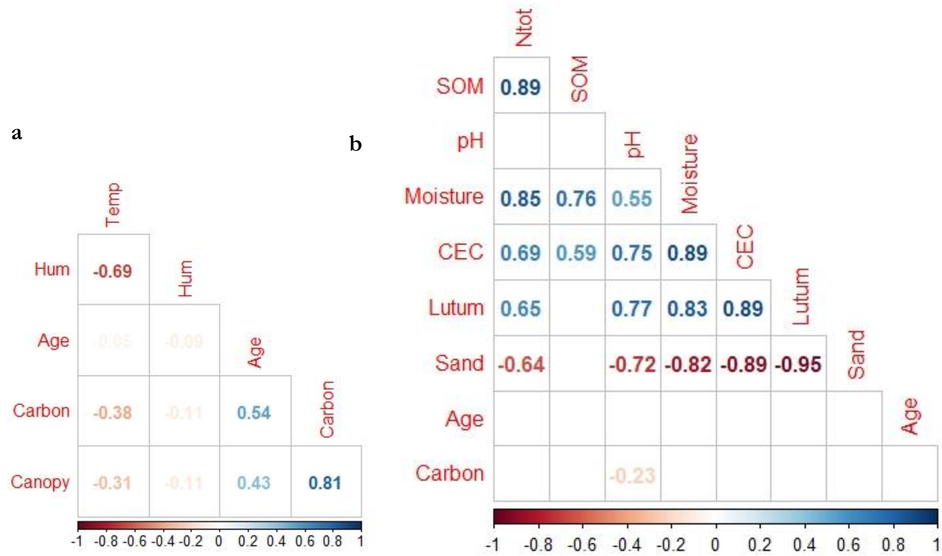


Figure A7: Correlation matrices of microclimate variables and soil variables separately. In both situations, the number of measuring points included in the research increased (from 12 to 31 and 18 respectively). **a)** Correlation matrix with aboveground carbon stock, canopy closure, difference in temperature and difference in humidity (both between inside and outside). Data was available per zone, n = 31. **b)** Correlation matrix with aboveground carbon stock and all seven soil variables. Data was available per food forest, n = 18. Carbon stock was now significantly correlating with soil acidity (pH). *Exported from RStudio.*

Table A5: p-values corresponding to Spearman's correlation tests including all variables. All green highlighted values show a significant correlation ($p < 0.05$), all orange highlighted correlations have a p-value between 0.05 and 0.10. All correlations are tested per food forest (n = 12).

	SOM	CEC	Moisture	Lutum	Sand	Age	pH	Carbon	Temp	Hum	Canopy	HeightMax
Ntot	9,8E-05	0,001247	4,55E-05	0,031283	0,030372	0,943418	0,320457	0,889373	0,493909	0,977393	0,988983	0,602153
SOM		0,031659	0,00665	0,338219	0,267608	0,764471	0,754667	0,648521	0,343238	0,708617	0,341459	0,08717
CEC			1,18E-06	5,95E-05	2,83E-05	0,934949	0,003768	0,442349	0,953484	0,491333	0,53265	0,827845
Moisture				0,000297	0,00055	0,862262	0,059749	0,670375	0,702447	0,799694	0,58892	0,948298
Lutum					7,81E-07	0,998041	0,006732	0,31242	0,813726	0,801219	0,137614	0,208867
Sand						0,767599	0,003174	0,432714	0,808569	0,592736	0,289918	0,461025
Age							0,370951	0,884081	0,752812	0,615684	0,261538	0,537972
pH								0,206494	0,518438	0,236651	0,565619	0,52036
Carbon									0,13721	0,166714	0,007525	0,009798
Temp										0,010857	0,565126	0,902241
Hum											0,204144	0,165866
Canopy												1,27E-05

Table A6: Correlation coefficients corresponding to Spearman's correlation tests including all variables. All correlations are tested per food forest (n = 12).

	SOM	CEC	Moisture	Lutum	Sand	Age	pH	Carbon	Temp	Hum	Canopy	HeightMax
Ntot	0,93007	0,811189	0,963	0,676073	-0,69231	0,145296	0,316909	0,251748	0,153846	-0,08392	-0,00356	0,192645
SOM		0,692308	0,874814	0,457758	-0,5035	0,232474	0,140849	0,559441	0,125874	-0,24476	0,316776	0,507882
CEC			0,899506	0,894388	-0,9021	0,217944	0,732413	0,076923	-0,02098	-0,23776	-0,18152	0,056042
Moisture				0,777976	-0,77604	0,194223	0,45826	0,194011	0,172846	-0,07055	-0,01436	0,143113
Lutum					-0,9789	0,087794	0,70922	-0,26067	0,049297	0,042255	-0,50899	-0,27514
Sand						-0,1344	-0,73945	0,188811	-0,00699	0,048951	0,44847	0,164624
Age							0,288988	0,537596	0,145296	-0,01816	0,325391	0,454846
pH								-0,19719	-0,10564	-0,30282	-0,35127	-0,10935
Carbon									-0,02797	-0,46853	0,832872	0,938705

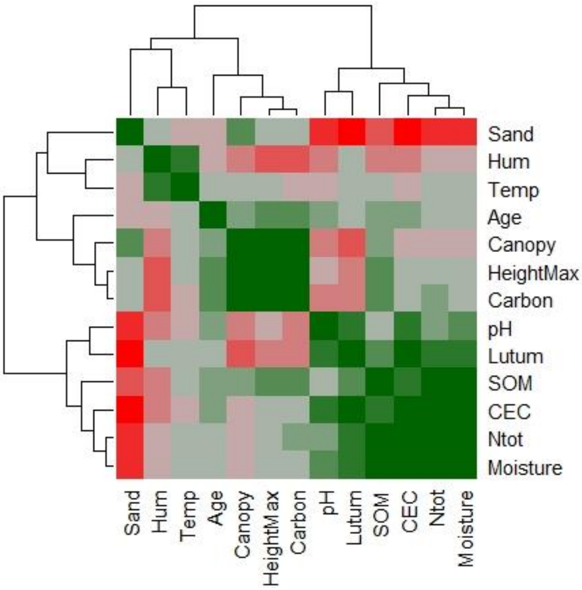


Figure A8: Heatmap of aboveground carbon stock, all microclimate variables, all soil variables and maximum height of tree layer. Two clusters can be distinguished: A: pH, Lutum, SOM, CEC, Ntot, Moisture. B: Age, Canopy, Heightmax, Carbon. *Exported from RStudio.*

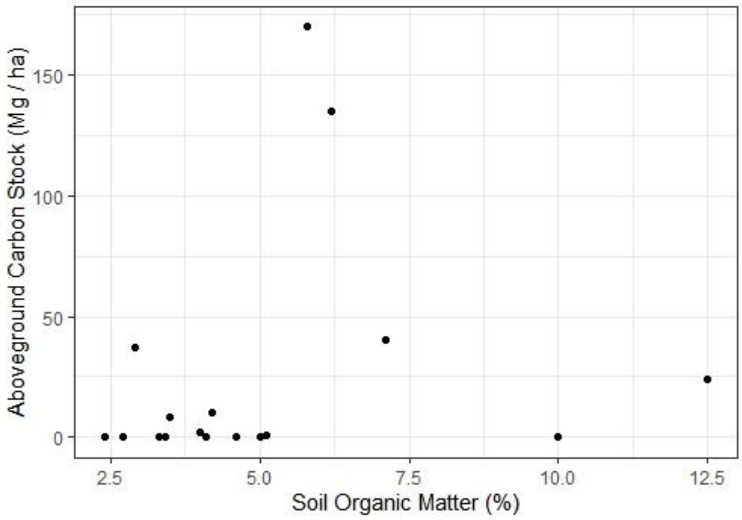


Figure A9: Food forest aboveground carbon stock versus soil organic matter. All plotted values represent food forests (n = 18). No significant regression was found ($p = 0.20$).

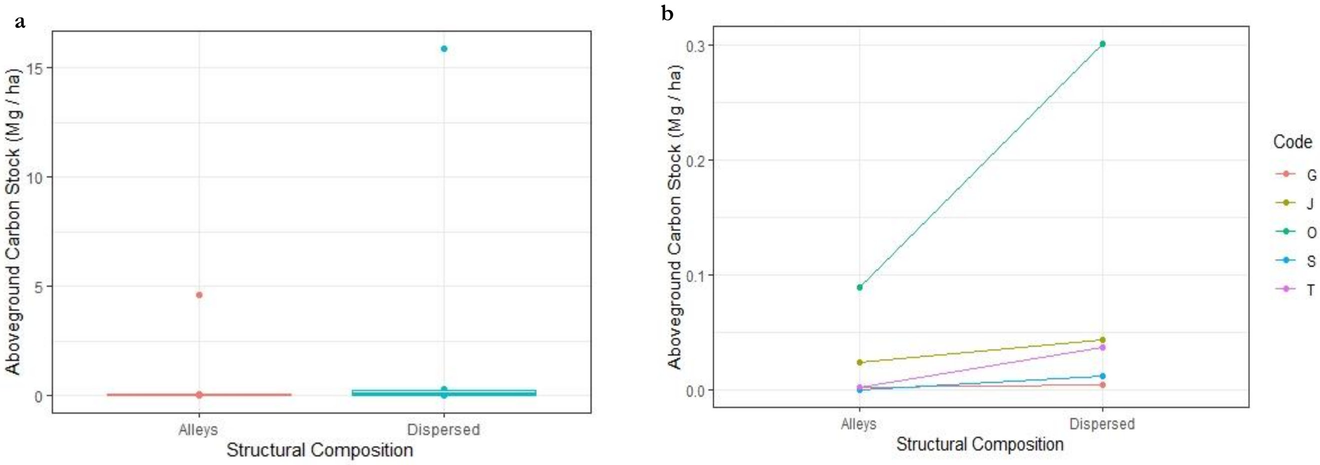


Figure A10: Supplementary plots on the relation between aboveground carbon stock and structural composition. Aboveground carbon stock is expressed in $Mg\ Co_2\ ha^{-1}$. **a)** Boxplot of the aboveground carbon stock per structural composition, for food forests with both compositions.n = 36. Dispersed food forests had a higher aboveground carbon stock than food forests with *alleys* ($p = 0.02$). **f)** Lineplot of food forests with both structural compositions. Strata means are shown. n = 12. Paired t-test showed significant higher carbon stock in dispersed food forests ($p = 0.02$). Food forest Ketelbroek is not shown, to improve visuality of the graph. *Exported from RStudio.*

Table A7: Outcomes of Tukey's Honest Significance Test for AGC stock per soil texture and former land use. a) Outcome of the post-hoc test for three categories of former land use. Significant differences were found between *forest* and *arable land* and between *forest* and *grassland* ($p < 0.05$). **b)** Outcome for the post-hoc test for three soil texture classes. Significant difference was found between sandy and clay soils ($p < 0.05$)

	Estimate	Std. Error	t value	Pr(>|t|)	
Forest - Arable Land == 0	64.811	10.187	6.362	<1e-05	***
Grassland - Arable Land == 0	2.000	9.387	0.213	0.975	
Grassland - Forest == 0	-62.811	10.331	-6.080	<1e-05	***

	Estimate	Std. Error	t value	Pr(>|t|)	
Loam - Clay == 0	17.91	11.81	1.516	0.28699	
Sand - Clay == 0	34.01	11.33	3.002	0.00922	**
Sand - Loam == 0	16.10	10.99	1.465	0.31112	

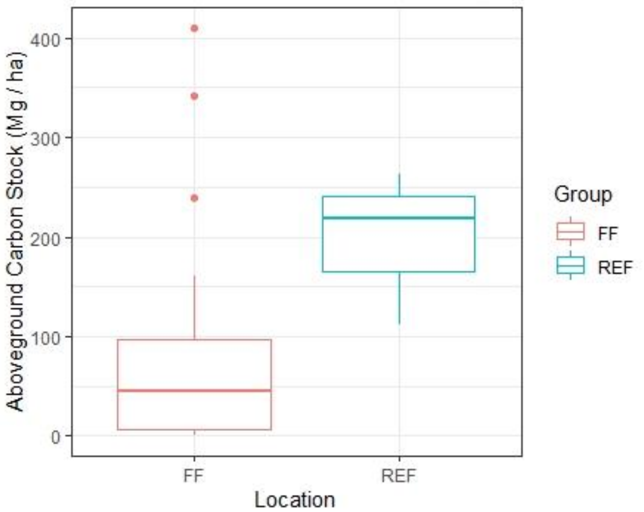


Figure A11: Boxplots of boveground carbon stock of food forests with FLU forest and reference forests. 35 sample plots of food forests and 3 sample plots of reference forests were included.. ACS was higher in reference forests, with 95% confidence interval = [24.48;228.0]; difference was significant ($p = 0.01$) based on Mann Whitney U test. *Exported from RStudio.*

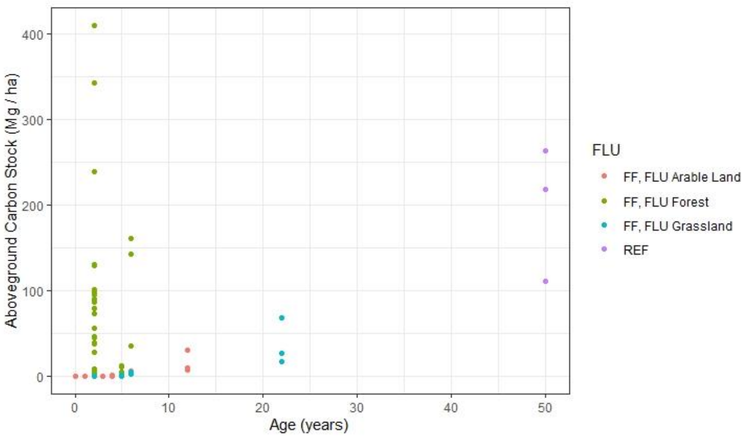


Figure A12: Aboveground carbon stock versus age of the (food) forest including reference forests. Data points are coloured by former land use. Food forests, n = 130), reference forests (n=3).

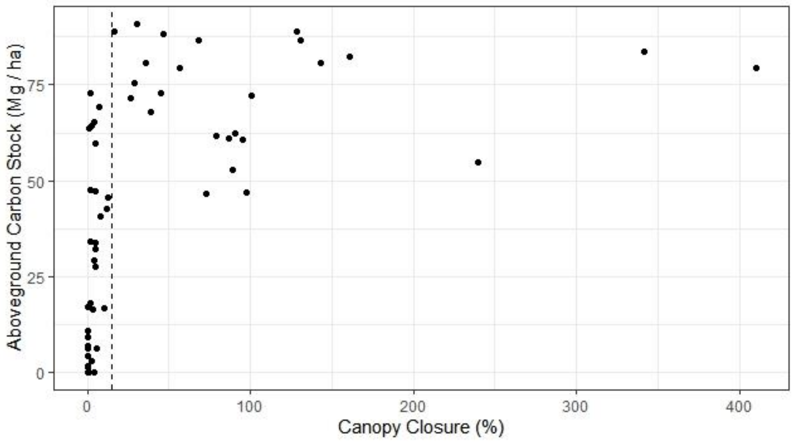


Figure A13: Aboveground carbon stock versus canopy closure. The turning point between a linear relation and the saturation of the curve is located at an AGC stock of ~ 15 Mg ha^{-1} which corresponds to an age of 7-8 years (for food forests without initial aboveground carbon stock).

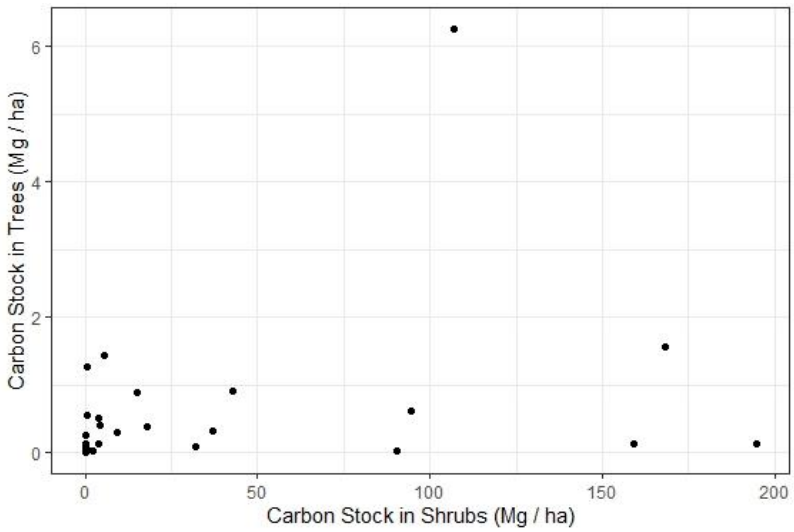


Figure A14: Carbon stock of shrubs versus trees on non-transformed scales. Scatter points representate sampling plots (n = 130). Aboveground carbon stock is expressed in Mg CO_2 ha^{-1}.

Table A8: The output of the *dredge* analyses for selecting the best explanatory generalised linear models concerning canopy closure (Canopy), maximum height (HT) and basal area (BA). Degrees of freedom (df) indicates model complexity; less degrees of freedom indicates a simpler model. Log likelihood (LogLik) indicates the explanatory capacity of the model; a lower LogLik indicates a better explaining model. Akaike Information Criterion (AICc) ranks models based on both explanatory capacity and complexity; the model with the lowest AICc is the best model. Delta AIC shows the difference in AICc between a model and the best scoring model; all models with a delta < 2 are selected.

Intercept	**Canopy**	**BA**	**HT**	**df**	**LogLik**	**AICc**	**Delta AIC**	**Adj. R^2**	**P**
-14.659	-0.601	0.0412	0.038	5	-346.20	703.29	0	0.99	2.2e-16
-17.124	*NA*	0.0389	0.019	4	-353.69	715.97	12.67	0.99	2.2e-16
-5.397	*NA*	0.0455	*NA*	3	-357.40	721.15	17.86	0.83	2.2e-16
-1.428	-0.173	0.0479	*NA*	4	-356.61	721.81	18.52	0.83	2.2e-16
-35.573	*NA*	*NA*	0.078	3	-393.25	792.85	89.55	0.55	2.2e-16
-35.064	-0.222	*NA*	0.087	4	-392.92	794.42	91.13	0.65	2.2e-16
-6.615	1.189	*NA*	*NA*	3	-410.35	827.05	123.75	0.28	8.5e-07
36.464	*NA*	*NA*	*NA*	2	-422.89	849.95	146.65	0.99	2.2e-16