Developing a methodology for the determination of homogeneous soil zones within long-term agroforestry experimental sites

Auteur : Mine, Raphaël
Promoteur(s) : Garré, Sarah; 12632
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : sciences et technologies de l'environnement, à finalité spécialisée
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/13887

LIÈGE université
Gembloux
Agro-Bio Tech

Developing a methodology for the determination of homogeneous soil zones within long-term agroforestry experimental sites

Raphaël Mine

Travail de fin d'études présenté en vue de l'obtention du diplôme de Master bioingénieur en sciences et technologies de l'environnement

Année académique 2021 - 2022

Co-Promoters : Pr. Sarah Garré (ULIEGE, ILVO), Dr. Bert Reubens (ILVO)

Le présent document n'engage que son auteur.

1. Dans ce cas, l'autorité académique est représentée par le(s) promoteur(s) membre du personnel(s) enseignant de GxABT

LIÈGE université
Gembloux
Agro-Bio Tech

Developing a methodology for the determination of homogeneous soil zones within long-term agroforestry experimental sites

Raphaël Mine

Travail de fin d'études présenté en vue de l'obtention du diplôme de Master bioingénieur en sciences et technologies de l'environnement

Année académique 2021 - 2022

Co-Promoters : Pr. Sarah Garré (ULIEGE, ILVO), Dr. Bert Reubens (ILVO)

Host Institution, scholarship and laboratories

This master thesis was carried out within the Research Institute for Agriculture, Fisheries and Food (ILVO) of Flanders in Ghent and with the help of the Erasmus Belgica program. Moreover, soil analyses have been partly conducted by a laboratory that is not part of ILVO : the agricultural analysis laboratory of the provincial center for agriculture and rurality in La Hulpe.

Acknowledgements

First of all, I would like to thank my co-promoter Pr. Sarah Garré for giving me the opportunity to do my master thesis in Ghent, at the ILVO. In this particular moment of Covid-19, it was a bit like travelling far away. Thanks also to my second co-promoter, Bert Reubens, for giving me the opportunity to work on this particular project and for introducing me to agroforestry. You both have always been understanding and available to help me.

Then, I would also like to thank Pr. Philippe De Smedt for always helping me and answering quickly when I had questions about electromagnetism.

I thank Pr. Catherine Charles, Pr. Gilles Colinet and Pr. Hugues Claessens who will take the time to read this master thesis and evaluate me afterwards.

I would like to thank my friends for always being there when I needed and for making these years in Gembloux Agro-Bio Tech incredible.

Finally, the most important, I would like to thank my parents who are the main reason for my success so far, who gave me the chance to study, who put everything in place for my studies to go well, who have always supported me and who know me better than anyone.

Abstract

On the one hand, the demand for locally produced nuts is increasing in Flanders while on the other hand there is a growing interest in agroforestry to help mitigate the environmental problems caused by conventional agriculture. In this context, a project coordinated by ILVO aims to develop agroforestry with walnut trees in Flanders and in the Netherlands. Part of their research focuses on late-budding varieties in order to delay competition for light with the crop and to avoid late frost damage. Seven long-term experimental trials on late-budding walnut trees were launched in 2021. The purpose of the present study was to identify zones with relatively homogeneous soil properties in four of the sites thanks to soil electromagnetic induction scan data clustering and soil analyses. The second objective was to analyse the impact of this zones on the performance of the walnut trees. Moreover, the methodology applied to this specific trial is intended to be applicable to other contexts. In each site, at least one zone with significant differences in terms of soil texture, water table level or water content has been identified. Although the current results are not conclusive as to the impact of these zones, differences in performance in some of them can be expected in the next few years, as the trees are still too young to draw any conclusions at this time. Finally, this method appears to have value in a broader context and can be applied to other agroforestry trials. Criticisms can however be made about the soil sampling method.

Résumé

D'une part, la demande pour des noix produites localement augmente en Flandre tandis que, d'autre part, l'agroforesterie suscite un intérêt croissant pour aider à mitiger les problèmes environnementaux causés par l'agriculture conventionnelle. Dans ce contexte, un projet coordonné par l'ILVO vise à développer l'agroforesterie avec des noyers en Flandre et aux Pays-Bas. Une partie de leurs recherches porte sur les variétés à débourrement tardif afin de retarder la compétition pour la lumière avec la culture et d'éviter les dommages causés par les gelées tardives. Sept essais expérimentaux à long terme sur des noyers à débourrement tardif ont été lancés en 2021. L'objectif de la présente étude était d'identifier des zones avec des propriétés de sol relativement homogènes dans quatre de ces sites grâce au clustering des données de scan du sol par induction électromagnétique et à des analyses de sol. Le deuxième objectif était d'analyser l'impact de ces zones sur la performance des noyers. De plus, la méthodologie appliquée à cet essai spécifique est destinée à être applicable à d'autres contextes. Dans chaque site, au moins une zone présentant des différences significatives en termes de texture du sol, de niveau de la nappe ou de teneur en eau a été identifiée. Bien que les résultats actuels ne soient pas concluants quant à l'impact de ces zones, on peut s'attendre à des différences de performance dans certaines d'entre elles au cours des prochaines années, les arbres étant encore trop jeunes pour tirer des conclusions à l'heure actuelle. Enfin, cette méthode semble avoir une valeur dans un contexte plus large et peut être appliquée à d'autres essais agroforestiers. Des critiques peuvent toutefois être formulées à l'égard de la méthode d'échantillonnage du sol.

Table of contents

List of figures

List of tables

Nomenclature

EMI = Electromagnetic induction
AF = Agroforestry
EC_a = Apparent electrical conductivity
HCP = Horizontal coplanar
PRP = Perpendicular
FDEM = Frequency Domain Electromagnetic Method
TDEM = Time Domain Electromagnetic Method
LUW = Laat Uitlopende Walnotenrassen
SFCM = Spatial fuzzy c-means

cLHS = conditioned Latin hypercube sampling
NDVI = Normalized difference vegetation index
ET_0 = Reference evapotranspiration
EU = European Union
XB = Xie-Beni
HAC = Hierarchical Agglomerative Clustering

1 Introduction

During the 20th century intensive agriculture allowed for high yields but this came at an environmental cost in terms of energy use, greenhouse gas emissions, water consumption and agrochemicals (Gomiero et al., 2011). Agriculture however also can play an important role to tackle climate change by promoting carbon sequestration in soils and biomass (European Commission, 2019). FAO 2021 report emphasises the importance of resilience in agri-food systems (FAO, 2021). A key element of this resilience is diversification and agroforestry is one example of climate-resilient practices. Agroforestry is described as "a collective name for land-use systems and technologies where woody perennials (trees and shrubs) are deliberately used on the same land management unit as agricultural crops and/or animals, in some form of spatial arrangement or temporal sequence" by Van Noordwijk et al. (2019). This type of agriculture might help reduce soil erosion and sequester carbon and therefore contribute to climate change adaptation and/or mitigation (IPCC, 2019).

In this context, the project "AGROFORESTRY 2025", coordinated by ILVO, was created with the aim of enabling Flanders to respond to the socio-economic challenges of today's agriculture and and to address the challenges of climate change resilience by increasing the applicability of agroforestry within the Flemish agricultural sector. The aim is to optimise the various facets of agroforestry systems in Flanders and to develop feasible financing and income models. The focus is therefore on the interaction between the development of innovative agroforestry systems at farm level on the one hand and innovative business arrangements on the other hand. Farmers therefore have an important place in the project, which is not restricted to the research community (ILVO et al., 2019).

This study takes place within the framework of this AGROFORESTRY 2025 project, more precisely within a part of this project about production and profit models for walnut trees in agroforestry. In addition to the positive effects of agroforestry, nut consumption, and particularly walnut consumption, is associated with a decreased risk of cardiovascular disease (Banel and Hu, 2009). Moreover, in the context of changing eating habits in Flanders, the demand for locally produced nuts is increasing and there is also potential for innovative markets for by-products such as nut oil (ILVO et al., 2019). In this project, several long-term experimental trials on walnut trees were launched in 2021. The purpose of these sites is, among other things, to examine different varieties of walnut trees and to compare known commercial varieties with experimental late budding varieties. The final purpose is to unlock the full potential agroforestry systems with walnut in Flanders by selecting varieties with the qualities of commercial ones in terms of productivity and growth, but also with the latest possible budding time to limit competition for light between the crop and the trees and to avoid late frost damage.

In a long-term experiment, it is essential to properly characterise the experimental sites. Differences in results between varieties must be analysed in considering a complete knowledge of the environment in which the trees grow. The purpose of this study is to characterise the soil conditions and heterogeneity in four of these sites and to identify the impact of these growth conditions on the tree performance. For this, we focus upon one research question : how does the soil heterogeneity in a field influence the performance of different walnut varieties ? To answer this question, this master thesis will be guided by two objectives : (1) the determination of zones with relatively homogeneous soil properties via soil electromagnetic induction scan data clustering and (2) the analysis of the impact of the previously determined zones on the performance of the walnut trees in terms of budding time and health.

In the future of the experiment, the different zones could lead to differences in terms of growth and productivity of the trees. The project is still young to actually be able to assess this impact, but the search for homogeneous zones is intended to provide a reading tool for potential future differences in performance between trees regardless of their variety. The methodology we present here, applied to this specific trial, is also intended to be applicable to other contexts.

2 State of the art

2.1 Agroforestry

2.1.1 Definition

As already mentioned, agroforestry (AF) is "a collective name for land-use systems and technologies where woody perennials (trees and shrubs) are deliberately used on the same land management unit as agricultural crops and/or animals, in some form of spatial arrangement or temporal sequence" (Van Noordwijk et al., 2019). In AF systems, there are both ecological and economic interactions between the different components (Van Noordwijk et al., 2019). These mixed systems are present in European landscapes since historical times (Eichhorn et al., 2006). The FAO proposes three main types of AF systems : silvoarable, silvopastoral and agrosilvopastoral (FAO, 2015). The first one consists of a combination of crops and trees, the second one mixes forestry and animals on pasture and finally the last combines trees, animals and crops. Agroforestry systems can be classified more precisely according to a large number of criteria : the structural basis (components, spatial arrangement of trees, vertical stratification and temporal arrangement of the different components), agro-ecological environment (environmental conditions and the ecological suitability of systems), socio-economical management level (level of inputs of management, intensity or scale of management and commercial goals) or function (major function or role of the system, mainly of the woody components) (NAIR, 1985; Herder et al., 2015). McAdam et al. (2009) proposes the table 1 based on NAIR (1985) which classifies AF systems.

Table 1 – Classification of agroforestry systems (McAdam et al., 2009).

Classification method	Example categories
(i) Components	**Silvoarable** : crops and trees
	Silvopastoral :pasture/animals and trees
	Agrosilvopastoral : crops, pasture/animals and trees
	Other : beekeeping and trees, aquaculture and trees
(ii) Predominant land use	**Primarily agriculture**
	Primarily forestry
(iii) Spatial and temporal arrangement	**Spatial**
	Mixed dense
	Mixed sparse
	Strip planting
	Boundary (e.g., trees on edges of plots and fields)
	In time
	Coincident, separate
(iv) Agroecological	Humid, arid, mountainous
(v) Socio-economic	Commercial, intermediate, subsistence
(vi) Function	**Productive function**
	Food, fodder, biofuel, wood, other products
	Habitat function
	Biodiversity
	Regulating functions
	Shelterbelt, soil and water conservation, shade
	Cultural function
	Recreation and landscape

2.1.2 Agroforestry in Europe and Flanders

AF systems in Europe leads to an improvement of ecosystem services especially in terms of soil erosion, soil fertility and biodiversity (Torralba et al., 2016). Moreover, carbon sequestration is enhanced and nutrient losses are decreased through the uptake of nutrients by the deep rooting trees (Rigueiro-Rodríguez et al., 2008). However, AF practices are still little implemented in Europe (Wezel et al., 2014). In Flanders in 2016, 55 % of farmers were not familiar with agroforestry and only 30 farmers were known to practice agroforestry

(Borremans et al., 2016). Nevertheless, in 2019, 50 farmers were known to practice it and it did not include the 21 subsidies applications for the 2019-2020 season (Reubens et al., 2019) so there is an increase in AF implementation in Flanders. A study of Graves et al. (2009) on 264 European farmers suggest that many farmers are open to the possibility implementing AF and that a good promotion and support of silvoarable agroforestry could lead to make it more common in Europe. However, even if considering AF is the first step in a possible implementation, it is not that simple and farmers tend to seek evidence before implementing an AF system such as experimental results or site visits (Graves et al., 2009).

A French study looked at farmers' reasons for adopting or rejecting agroforestry (Mary et al., 1998). The main technical argument against AF is in general that mechanisation is difficult and dangerous on some plots (small or sloping). In addition, the lack of machinery is a factor against AF. Also, they fear the impact of competition between the different components of the system when soils are not very fertile or when water is scarce. On the contrary, the arguments in favour of AF are the intensification of agricultural production, the savings on orchard maintenance costs and the improvement of soil fertility or the help in controlling pests and weeds along the tree rows (Mary et al., 1998). In Belgium, nowadays, farmers mainly do AF for diversification purposes and to make sense of what they are doing from an ecosystem services perspective according to Olivier Baudry, forester expert and secretary and referent of ongoing projects for the association for the promotion of agroforestry in Wallonia and Brussels (AWAF). It is confirmed by the study of Graves et al. (2009) who explains that, in Northern Europe and contrary to the Mediterranean areas, farmers consider that the main advantages of silvoarable systems are environmental.

There are not only technical arguments, but also economic and institutional bottlenecks that explain the low number of AF systems in Northern Europe and more precisely in Belgium. First, the uncertainty related to the profitability of AF systems is one of the main obstacles to their adoption by farmers (Borremans et al., 2018). Nowadays, the main financing sources for AF system are the subsidy program and greening measures and there is still a need to the development of special markets for AF products (Borremans et al., 2018). An other issue concerns the the ownership of the land. On the one hand, in Flanders in 2016, 64 % of the farms were leased while, in the other hand, farmers only tend to plant trees on farmland they own partly because of the tenancy law in Flanders which imposes a request for permission to plant trees to the owners (Borremans et al., 2016). Another legal bottleneck is the uncertainty about whether trees can be harvested because of their value for biodiversity, nature, and society. There are indeed still potentially conflicting rules and decrees that apply to AF systems in Flanders (Borremans et al., 2016). Finally, there is a lack of trust in the government by the farmers who consider that the government does not implement a steady policy Borremans et al. (2018). It is therefore clear that a major barrier to AF implementation in Flanders is the lack of economic and institutional incentives for farmers (Reubens et al., 2019).

2.1.3 Interaction of components

The productivity of an AF system is the result of positive and negative interactions between the different components (Jose et al., 2004). These interactions can be divided into aboveground and belowground interactions. Belowground interactions mainly include competition, complementarity or facilitation in the search for water and/or nutrients but also allelopathy. Among the aboveground interactions we can mention the microclimatic modifications (wind, temperature, air humidity), the shading of the trees on the crop and the density and diversity of insects that can be natural enemies of pests (Ong et al., 1991; Jose et al., 2004).

In Northern Europe, light is probably the main constraint for AF system productivity (Eichhorn et al., 2006). Indeed, in the absence of water or nutrient stress, the biomass growth of the crop depends on the amount of photosynthetically active radiation intercepted by the canopy (Monteith et al., 1977) and therefore, shading of trees generally results in reduced growth and yield of crops because they induce a microclimate that causes reduced solar radiation (Martsolf, 1966 cited by (Lin et al., 1999)). Therefore, the budding time of trees in an AF system is very important to consider. The later the budding time, the later the competition for light between the trees and the crop begins. Light competition can also be minimised by pruning the trees, which also encourages air circulation in the system, reducing the risk of disease, and allows

the production of high value knot-free wood (Pantera et al., 2018), and through smart design (orientation of the tree rows, distance between the tree rows, adapting crop species and variety choice to partial shade conditions).

2.2 Walnut tree

The common walnut tree (*Jugans regia* L.) is very well adapted to AF systems and offers many advantages (Consortium Agroforestry Vlaanderen, 2021). Firstly, the competition for light is relatively low with agricultural crops. The crown lets in a reasonable amount of light and the tree has leaves only for a short period from mid-May to October (Germain et al., 1999; Consortium Agroforestry Vlaanderen, 2021). As far as the soil quality is concerned, the leaves of the walnut trees have a good litter quality, which has a positive effect on soil fertility (Consortium Agroforestry Vlaanderen, 2021). Moreover, harvesting nuts provides income in the relatively short term compared to timber production. Finally, market demand for walnuts is increasing due to an increased awareness of the positive impact of nuts on health (Bemah et al., 2012). A drawback could be possible due to the production of an allelopathic compound, juglone. Juglone has an inhibitory effect on the growth of many types of plants. However, it is mainly produced by the black walnut (*Juglans nigra* L.) while the common walnut produces such limited amounts of juglone that toxicity to other plants is rarely observed (Consortium Agroforestry Vlaanderen, 2021).

2.2.1 Floral biology of walnut trees

The common walnut tree inflorescences are unisexual (Germain et al., 1999) (except for some Asian varieties). Both male and female inflorescences are present on the same tree, making the walnut a monoecious species. Male inflorescences are composed of 100 to 160 flowers to form a catkin while the female inflorescences are single or grouped by two or 3 (Germain et al., 1999).

Jugans Regia L. is a dichogamic species which means that male and female inflorescences are delayed in time. There are both protandrous and protogynous varieties (Germain et al., 1999). Budding-time of female inflorescence, which is concomitant with the leafing, varies a lot within a species depending on the climate and between species. This budding period is important to consider, even when walnut trees are not in an AF system, because spring frosts reduce the production potential of trees when budding has already occurred (Germain et al., 1999). Late budding varieties are therefore interesting on the one hand to avoid losses of productivity of the tree but also to reduce the competition for light with the crop as much as possible.

2.2.2 Walnut tree multiplication

Common walnut trees can be multiplied by sexual reproduction, grafting or micropropagation (Germain et al., 1999). Concerning the sexual reproduction, sowing is still used in some countries but leads to heterogeneous orchards (Germain et al., 1999). The reason is that sexual reproduction does not allow for identical reproduction of desired characteristics. However these seedlings can be used as rootstocks (Germain et al., 1999). Walnut trees can be grafted in a nursery (most of the time) or directly on field (Germain et al., 1999). Grafting in a nursery has advantages in terms of the homogeneity of the grafted plants. However, in Belgium, there are few nurseries performing grafting, particularly because the technique is quite specific and the success rate is limited [pers. comm. Bert Reubens and Olivier Baudry].

2.2.3 Soil

The good development of the root system is very important in order to explore a sufficient volume of soil and to meet the water and nutritive needs of the walnut tree. This development is mainly guided by the soil physical properties (Germain et al., 1999).

First, to meet the water needs in case of drought, the roots need to go deep and therefore a relatively deep soil is necessary for walnut (Germain et al., 1999). Even if most of the roots are situated in superficial layers of the soil and that the density of deeper roots is far lower, these deeper roots have a more important role

in uptake of water than indicated by their density alone and can therefore maintain a good water supply during droughts (Stone and Kalisz, 1991; Germain et al., 1999).

As far as texture is concerned, figure 1 shows in green the texture classes favorable to walnut trees. Moreover, soils with impermeable layers, where water is not sufficiently drained, should be avoided for the good development of walnut trees because the roots are quite sensitive to lack of air (Wertheim, 1981).

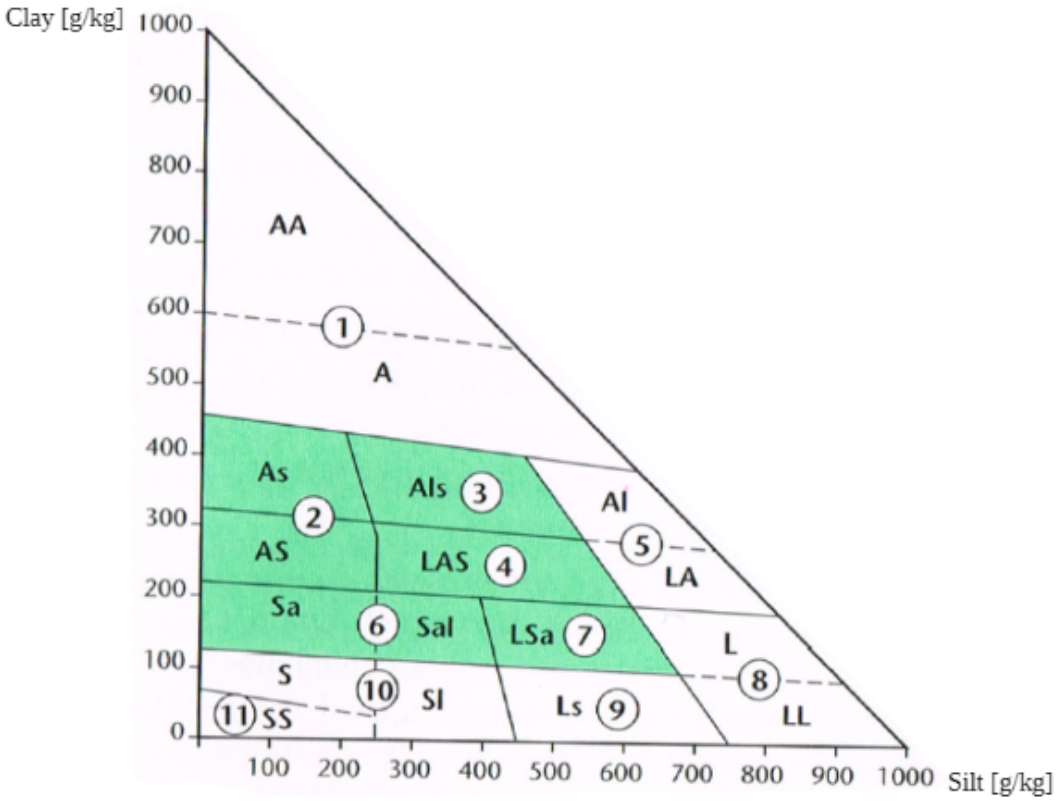


Figure 1 – Textural triangle for walnut trees (G.E.P.P.A. (Groupe d'études des problèmes de pédologie appliquée) cited by (Germain et al., 1999)).

With :

1 : very fine and clayey texture : AA and A
2 : fine, sandy-clay texture : As and AS
3 : fine texture of silty-sandy clay : AIS
4 : fine silty-clay-sandy texture : LAS
5 : clayey silt texture : AI and La
6 : medium sandy texture : Sa and SaI
7 : medium clayey silt texture : LSa
8 : medium silty texture : L and LL
9 : coarse texture of sandy silt : Ls
10 : coarse sandy texture : S and SI
11 : very coarse sandy texture : SS

Soil structure is also important for the roots to facilitate the penetration. The compaction caused by agricultural machinery degrades this structure by increasing the density of the soil (Germain et al., 1999).

For the nutritive status, a good production and feeding of walnut trees, especially with lateral fruiting (most of commercial varieties), the soil must be fertile (Germain et al., 1999). Indeed, although walnut trees are able to absorb some nutrients directly with their leaves, the majority of nutrient requirements are met from the soil solution (Germain et al., 1999).

Regarding the pH, the suitable pH H_2O is within the range of 6 to 7.5 (Germain et al., 1999) but there is a tolerance between 5.5 and 8.5 (Wertheim, 1981). The reason is that pH influences the availability of chemical elements, as shown in Figure 2, and should therefore ideally be kept between those values.

Finally, as far as organic status is concerned, the optimal C/N ratio is around 10 for walnut trees (Germain et al., 1999).

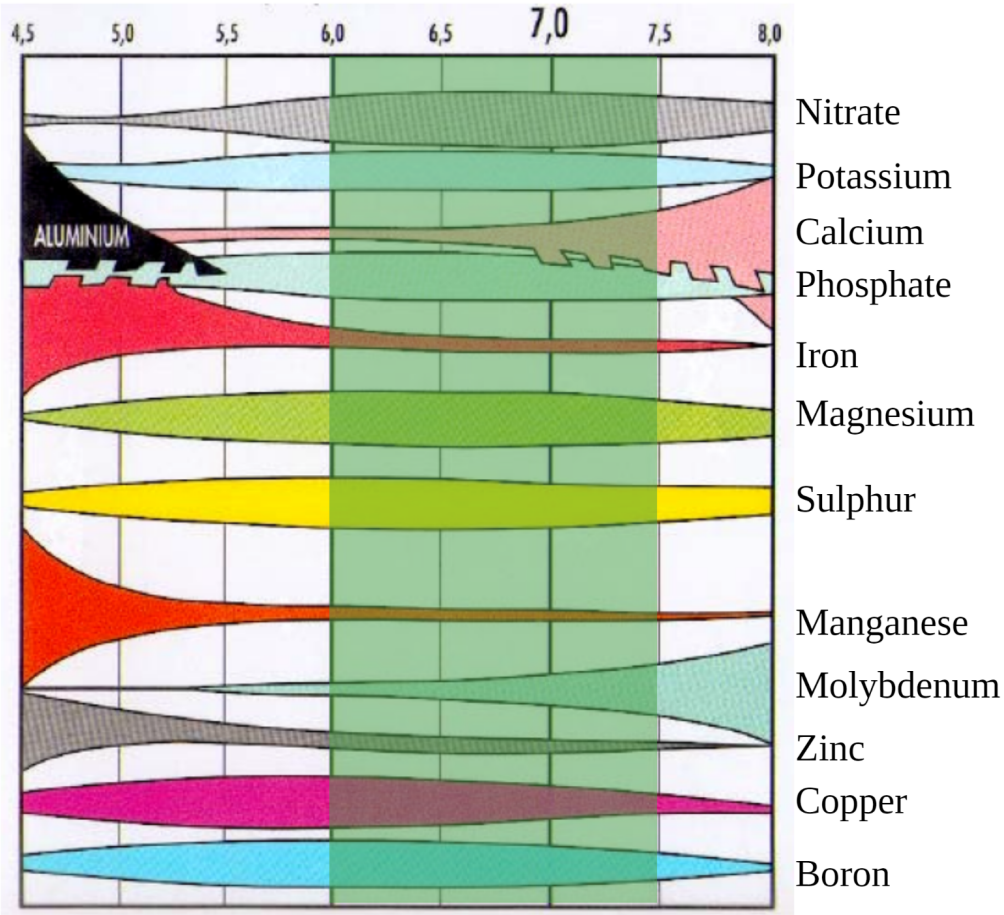


Figure 2 – Diagram of assimilability of elements according to pH (Emil TRUOG's diagram) (Merelle, 1998; Genot et al., 2007).

To conclude this section about the soil it is important to note that globally the common walnut tree is not very demanding regarding the type of soil, but the most important elements for it to grow best are a deeply rooted and well aerated soil, with a sufficient supply of water (Germain et al., 1999) but not lack of air and rather rich in organic matter (Wertheim, 1981; Consortium Agroforestry Vlaanderen, 2021).

2.2.4 Overview of walnut market

Two types of market for walnuts exist : in-shell walnuts and shelled walnuts. The market of shelled walnuts is the most important (Germain et al., 1999). In addition to the shelled and in-shell walnuts market, pressing walnut oil is the typical value addition to the walnut value chain (Șvarț, 2018).

The database of the FAO shows that the world walnut production (in-shell) in 2019 was around 4.5 millions of tons (Table 2). China is the leading producer by far, followed by the United States. In season 2020/2021, China production was accounting for 47 % of the world share (INC - International Nut and Dried Fruit Council, 2021).

Table 2 – World walnut production (in-shell) for countries with a production over 20 000 tons from 2009 to 2019 (in thousands of tons) (source : FAO).

	2009	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
World	2475	2767	3196	3655	3026	3354	3878	4063	4201	4347	4498
China	979	1284	1656	2021	1454	1535	1942	2114	2250	2386	2522
USA	396	457	418	497	446	518	550	625	572	616	592
Iran	280	268	280	284	223	403	420	349	394	304	321
Turkey	177	178	183	203	212	181	190	195	210	215	225
Mexico	115	77	96	111	107	126	123	142	147	160	171
Ukraine	84	87	113	97	116	103	115	108	109	127	126
Chile	26	32	35	40	43	70	90	90	100	110	123
Uzbekistan	14	14	34	30	40	44	52	53	47	48	51
Romania	38	34	35	31	32	32	33	34	46	54	50
France	42	32	38	36	36	35	42	40	33	38	35
India	36	38	36	40	36	43	35	33	/	/	/
Greece	22	23	23	23	24	25	26	29	36	32	31
Egypt	22	21	22	24	24	24	24	24	24	24	24
Serbia	25	21	24	13	18	16	17	14	12	10	9

Within the European Union (EU), 65 % of domestic consumption is supplemented by imports even though the EU produces its own nuts (Șvarț, 2018). In fact, there is an increase in consumption (Bemah et al., 2012) but this is accompanied by an increase in imports rather than an increase in production (Șvarț, 2018). In 2015, the United States was the leading supplier of walnuts to the EU by far, while Moldova and Chile were in second and third place respectively (Șvarț, 2018). France and Romania are the most important producers within the EU while Germany is the main importer (Șvarț, 2018). In Belgium, the production is still very small and it is estimated that only 1 to 2 % of the nuts in Belgian stores are of Belgian origin, according to Guy Lejeune, a pioneer in the walnut production in Wallonia. The Belgian walnuts market is mainly based on direct sales into the farm and without intermediaries or in cooperatives (pers. comm. Olivier Baudry). Some walnut producers sell walnut derived products in addition of simply walnuts. For example, Guy Lejeune who sells walnut wine and walnut oil (pers. comm. Olivier Baudry). For Olivier Baudry, the nut sector should be encouraged in Belgium for several reasons. First of all, there is a risk of drought issues in France due to climate change and therefore less production. Secondly, walnut production offers a possibility of diversification and adaptation to climate change for farmers. Finally, local products are in vogue and importing almost 100 % of a product means being dependent on other countries. However, nowadays, there are still some strong challenges in terms of economic viability for the nut production in Belgium (pers. comm. Olivier Baudry).

As far as organic walnut production is concerned, the market is still quite small, although interest in organic food has been growing for two decades. Moldova is the leading exporter of organic walnut kernels to the EU, followed by the US and France (Șvarț, 2018).

2.3 Soil electromagnetic induction scan

As previously mentioned, although walnut trees are not overly sensitive to soil types, some soil characteristics are important to consider, including texture, structure and moisture content. It is therefore important to assess the soil conditions to fully understand the fields where the trees are planted within the framework of a project like "AGROFORESTRY 2025". One way to do this is the electromagnetic induction (EMI) scan that measures the apparent electrical conductivity (EC_a) of a subsurface soil volume (Allred et al., 2008), EC_a being nowadays the most valuable geophysical measurement to characterize the soil spatial heterogeneity of a field (Corwin, 2008). Electrical conductivity is called "apparent" in this case because it is an average EC_a for a volume of soil (Doolittle and Brevik, 2014).

Frequency electromagnetic induction methods have been used for a long time now for geophysical characterisation purposes (McNeill, 1980). Two modalities exist : time domain (TDEM) and frequency domain

(FDEM). Today, FDEM is often and efficiently used to study 3D spatial and temporal variations in soil (Boaga, 2017). The success of this method has several explanations. First, the sensors used are light and do not require contact with the ground, which facilitates the work of technicians and reduces the time of investigation (Calamita et al., 2015). In addition, depending on the coil distance, EMI sensors measure through soil thicknesses that can be up to 100 m for TDEM and several meters for FDEM (Robinson et al., 2008b). Then, concerning soil moisture measurements, local measurements at specific points using soil moisture sensors have a limited spatial resolution. Indeed, a manual survey requires a lot of work and it is cost-intensive when talking about network installations. On the contrary, the EMI method provides information with a much better spatial resolution, even for large fields (Vereecken et al., 2014). Finally, soil electrical conductivity, which is measured during EMI scan, depends on a combination of mineralogy, clay content, soil moisture, pore water electrical conductivity (i.e. salinity) and soil temperature (McNeill, 1980; Vereecken et al., 2014; Calamita et al., 2015; Corwin and Scudiero, 2019) and can therefore be used to map the soil properties.

To measure the EC_a of a soil volume, EMI sensor transmit an electromagnetic field in this soil. In this way, electrical currents are induced in the soil, which in turn generate an electromagnetic field. This secondary electromagnetic field is read by the sensor's receiver. The secondary field is proportional to the electrical current for most soil types and it is so possible to calculate the apparent EC_a of the explored volume of soil under certain conditions (Doolittle and Brevik, 2014). The following part deals more precisely with the physical principles and formulas underlying the EMI method, which are adapted from Boaga (2017). First, the Biot-Savar law (equation 1) states that "a uniform electrical current produces a magnetic field in the vacuum, whose magnitude, B, depends on the current strength, I, and on the radius vector, r, between the current line and the measurement point".

$$B = \frac{\mu_0}{4\pi} I \int \frac{dl \times r}{r^2} \tag{1}$$

With :

μ_0 : the magnetic permeability of the vacuum, $4\pi \times 10^{-7}$ kg m A^{-2} s^{-2}.
dl : the unit vector along the current line.

FDEM instruments generate an alternating current to induce an alternating magnetic field in the soil that will also induces an electromotive force (e.m.f.). This force can be calculated thanks to the Faraday's law (equation 2).

$$\text{e.m.f.} \ = -\frac{d\phi}{dt} = -2\pi i f \phi \tag{2}$$

With :

ϕ : the magnetic flux [kg m^2 A^{-1} s^{-2}].
f : the frequency [Hz].
$i = \sqrt{-1}$: imaginary unit.

The electromotive force induced by the primary magnetic field (Hp) is out-of-phase comparing to the magnetic flux and produces secondary electrical currents. These currents cause, in turn, a secondary magnetic field (Hs). The difference between Hp and Hs depends on the soil properties and the geometry of the coils (spacing between coils and orientation and distance from the soil surface) (Hendrickx and Kachanoski, 2002; Boaga, 2017). When working in 'low induction number' conditions (i.e. $2\pi f \ll 2/\mu_0$ EC s^2), EC_a can be calculated as in equation 3.

$$EC_a = \frac{4}{\omega \mu s^2} \frac{Hs}{Hp} \tag{3}$$

With :

s : the inter-coil spacing [m].
ω : the angular frequency [rad s^{-1}].

The investigated depth depends on the generated electro-magnetic forces (Boaga, 2017) and therefore depends on the parameters of the measuring device as the coil orientation and spacing (Doolittle and Brevik, 2014).

Before using the measured data, it is important to ensure reliable data. It is therefore necessary to go through a dedicated data processing, taking into account specific application and instrument setup (Hanssens et al., 2020). First, a data calibration and filtering step is necessary because data errors can strongly impact the results and must therefore be taken into account (Minsley et al., 2012). Two types of errors are possible : random errors and systematic errors. Systematic errors can be due to incorrect instrument calibration, drift, and improper data levelling (Minsley et al., 2012). Drift corresponds to a response at one given location that varies over time, despite no appreciable changes above or underneath the surface (Delefortrie et al., 2014). In addition to this calibration, FDEM data must be accurately georeferenced by considering spatial offsets between the GPS antenna and FDEM instruments or offsets induced by a time lag between them (Hanssens et al., 2020).

EMI has lots of application in soil surveys. It is used to assess soil salinity (Corwin and Scudiero, 2019), subsurface water movement and soluble salts, soil water content, soil texture, clay content, soil compaction, CEC, pH, exchangeable Ca and Mg, carbonates, soil organic carbon, leaching rates of solutes, herbicide, nitrogen, differences in terms of lithology and mineralogy, (Doolittle and Brevik, 2014) or plant community spatial patterns (Robinson et al., 2008a). Even if it is less related to soil science, EMI is also used for the detection of buried utilities and services and the detection and tracing of archaeological remains (Calamita et al., 2015).

2.4 Spatial data clustering

Prior to looking at spatial data, it is important to define what data clustering is. It is "the unsupervised classification of patterns (observations, data items, or feature vectors) into groups (clusters)" (Jain et al., 1999). Data set is therefore split and organized into several clusters based on the internal homogeneity and the external separation (Jain et al., 1999; Xu and Wunsch, 2005). Clustering is called "unsupervised" because the definitions of the classes and its number are not known in advance (McClean, 2003). On the contrary, in a supervised classification, the first step is to select training samples with known classes to use them to guide the computer to identify the features of each class (Wu, 2018).

The most known clustering algorithms are Hierarchical Agglomerative Clustering (HAC) and k-means clustering (Gelb and Apparicio, 2021).

HAC basically consists in successively aggregating the observations according to their degree of similarity until a single class is obtained (Gelb and Apparicio, 2021). First, it is necessary to calculate the proximity matrix which includes the distances between each pair of observations. Then, most similar pair of clusters, based on the previously calculated matrix, are merged to form a new cluster. Finally, the algorithm stops when all patterns are into one single cluster (Jain et al., 1999). For n individuals, the first partition has therefore n-1 classes, then n-2 classes and so on until a partition gathering all the patterns (Gelb and Apparicio, 2021). This hierarchy of partitions can be represented by a tree-like nested structure partition of the data set, called dendrogram (Jain et al., 1999; Xu and Wunsch, 2005). By looking at this dendrogram, the appropriate number of clusters can be decided.

Concerning k-means clustering, the grouping in k groups is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid (Nagpal et al., 2013). In the case of this method, the number of groups is decided in advance, it is a partitional approach and not a hierarchical one as for HAC (Jain et al., 1999). The objective function (function to minimize) of k-means is given in equation 4 (Nagpal et al., 2013).

$$\mathrm{E} = \sum ||\mathrm{X}_i - \mathrm{m_i}||^2 \tag{4}$$

With :

E : sum of square error for all objects in the data.
X_i : a pattern in a cluster.
m_i : the middle of a cluster k_i.

In other words, the first step of k-means clustering is to choose a number of clusters, k. Then, k defined points are randomly-chosen inside the hypervolume containing the pattern set, they are the first k cluster centers. Afterwards, each patterns is assigned to the closest cluster center. The next step is to recompute the cluster centers using the current cluster memberships. If a convergence criterion is not met, the algorithm starts again at the data point cluster assignment step (Jain et al., 1999).

HAC and k-means clustering are said to be hard clustering because a point can belong to one cluster only (Zass and Shashua, 2005). On the other hand, there are so-called soft or fuzzy clustering methods. In a fuzzy classification each pattern has a variable degree of membership in each of the output clusters (Jain et al., 1999). The fuzzy c-means (FCM) method will be developped here. FCM method is therefore a soft and unsupervised clustering technique introduced by (Bezdek, 1981). In a fuzzy c-partition, the sample points get a degree of membership to each group that varies between zero and one (Bezdek et al., 1984). The higher a membership value is for a sample, the greater is the certainty in the cluster assignment (Stetco et al., 2015; von Hebel et al., 2021). This algorithm has been used a lot in soil clustering (Goktepe et al., 2005; Hanesch et al., 2001; von Hebel et al., 2021; Schröter et al., 2017; Martinez et al., 2009). The use of a fuzzy classification is relevant in soil science because of "the complexity of soil variation at all scales and the resultant uncertainty in soil mapping and spatial interpolation" (McBratney and Odeh, 1997).

However, even if FCM is used a lot with soil data, it doesn't take spatial information into account. Not considering this information can lead to noise, some noisy points being classified in one group while all surrounding points are classified in another (Chuang et al., 2006). The use of spatial information in FCM has therefore been developed to deal with this issue. The first scientific field to have studied SFCM is the field of magnetic resonance imaging (MRI) of the brain (Chuang et al., 2006; Cai et al., 2007).
(Gelb and Apparicio, 2021) propose to use SFCM in geographical field. To include spatial information, a parameter α is added to FCM. It ranges between 0 and $+\infty$ represents the weight of the spatial dimension. Furthermore, users of *geocmeans* have to define a binary (0,1) spatial matrix which represents the neighbours spatial weights. The matrix must have odd dimensions and the center of the matrix represent the focal point.

The formulas behind SFCM are given below (Gelb and Apparicio, 2021). Equation 5 is the objective function of the SFCM.

$$J_m = \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^m \left\|x_k - v_i\right\|^2 + \alpha \sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^m \left\|\underline{x_k} - v_i\right\|^2 \tag{5}$$

With :

v_i : the center of gravity of the group i.
c : the number of groups.
x_k : the value of the kth pixel .

$\underline{x_k}$: a means of neighboring pixels lying within the defined window around x_k.
m : the fuzziness degree parameter.
α : a parameter used to control the effect of the neighbors term.
N : the number of observations.
u : the belonging matrix (degree of membership for each cluster). u_{ik} is therefore the probability for the observation k to belong to the group i. By definition, each sample point x_k satisfies the constraint that $\sum_{i=1}^{c} u_{ik} = 1$ (Cai et al., 2007).

u_{ik} and v_i can be calculated as in equation 6 and 7.

$$u_{ik} = \frac{\left(\|x_k - v_i\|^2 + \alpha \|\underline{x_k} - v_i\|^2\right)^{-1/(m-1)}}{\sum_{j=1}^{c} \left(\|x_k - v_j\|^2 + \alpha \|x_k^{-v_j\|^2)^{-1/(m-1)}}\right.} \tag{6}$$

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^m (x_k + \alpha x_k)}{(1+\alpha) \sum_{k=1}^{N} u_{ik}^m} \tag{7}$$

3 Material and method

3.1 Experimental sites

This project worked on 4 selected sites in a long-term experimental network on late budding walnut varieties. They are located in Merelbeke (East Flanders), Herent (Flemish Brabant) Galmaarden (Flemish Brabant), and Hulste (West Flanders) (figure 3). They are respectively coded as "LUW1", "LUW2", "LUW3" and "LUW4". "LUW" is for "Laat Uitlopende Walnotenrassen" which means late-budding walnut varieties in Dutch.

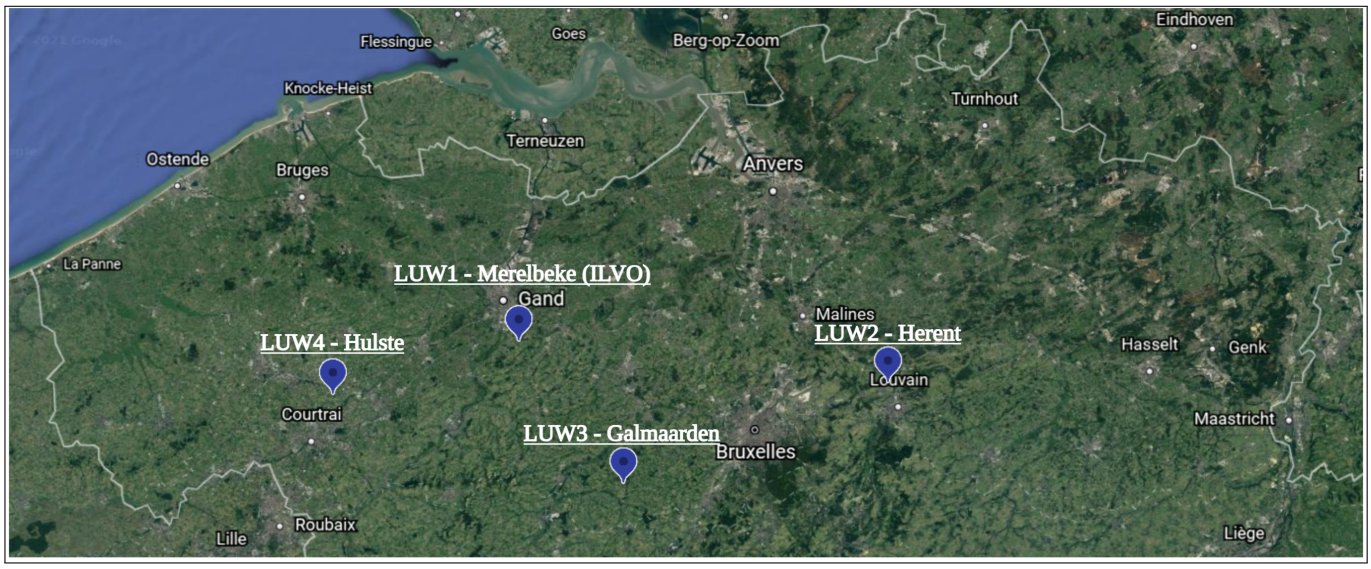


Figure 3 – Experimental sites location.

3.1.1 LUW1 - ILVO

LUW1 site is an 1.1 ha cropland plot presented in figure 4. For the moment only 0.7 ha are devoted to the agroforestry experiment but it could evolve in the future. This site is managed by ILVO.

Soil type of this site is a Lba0 over its entire extent. It means that the soil is a sandy loam with good drainage, B textural horizon (B_t) which means that there is a clay accumulation in depth and finally the A horizon is more than 40 cm thick.

The plot is located on a plateau and has very low slopes. The difference is only a little over 1.4 m between the highest and lowest point.

The AF experimental design of ILVO site is presented in figure 5. The code for the different varieties is explain in the point 3.2 (table 3). The plot has three walnut tree rows of 20, 21 and 22 walnuts trees. The 63 trees are planted in such a way that each variety is present at least once in every tree row. Moreover, the LUW varieties and commercial varieties are separated from each other, resulting in four blocks. While currently the focus of the experiment is on the monitoring of the trees, this design will allow to assess the impact of the LUW trees versus the commercial trees on crop performance in the alleys between the tree rows in a later stage. Tree rows are 24 m apart and trees in the same row are 5 m apart.

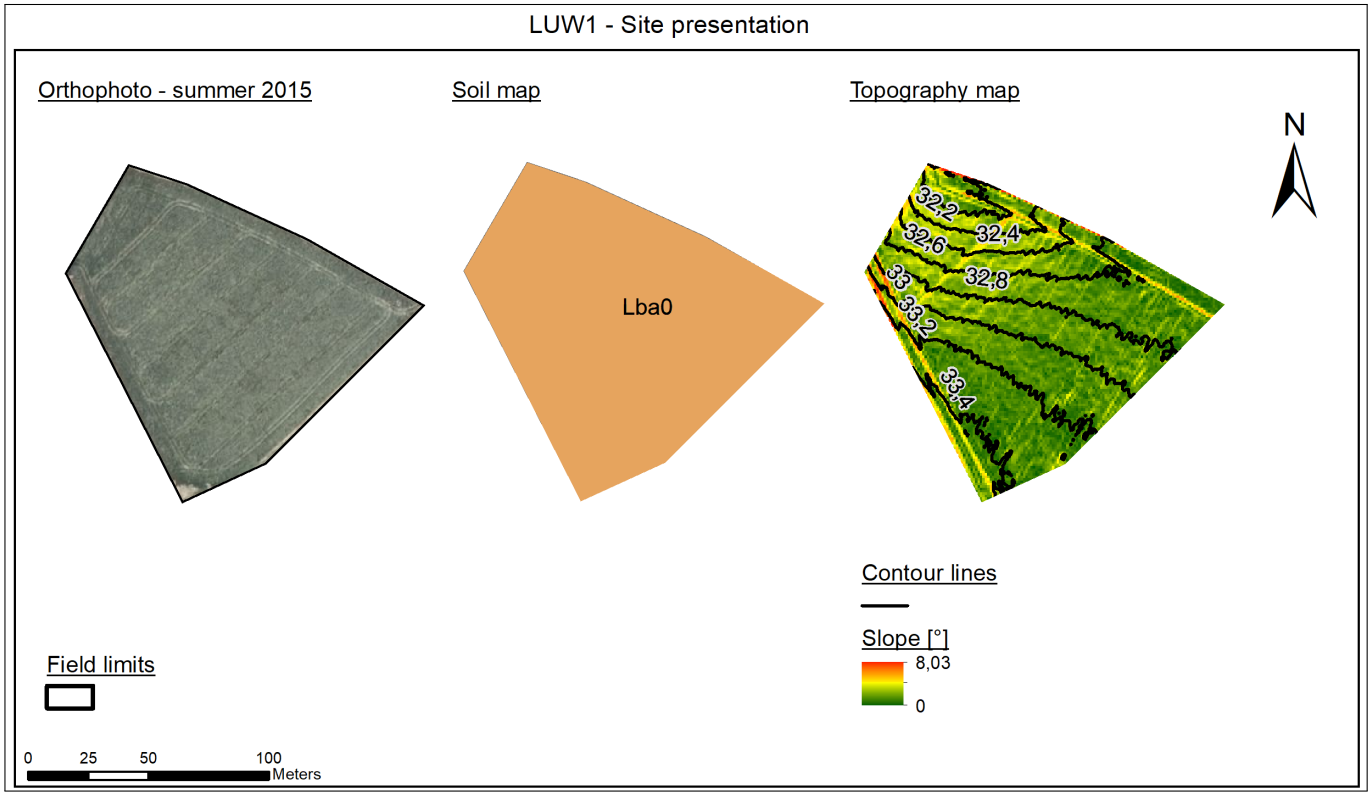


Figure 4 – LUW1 - Orthophoto, soil map and topography (1 m x 1 m).

All data layers for the presentation maps of the four sites come from the *Geopunt Vlaanderen* website.

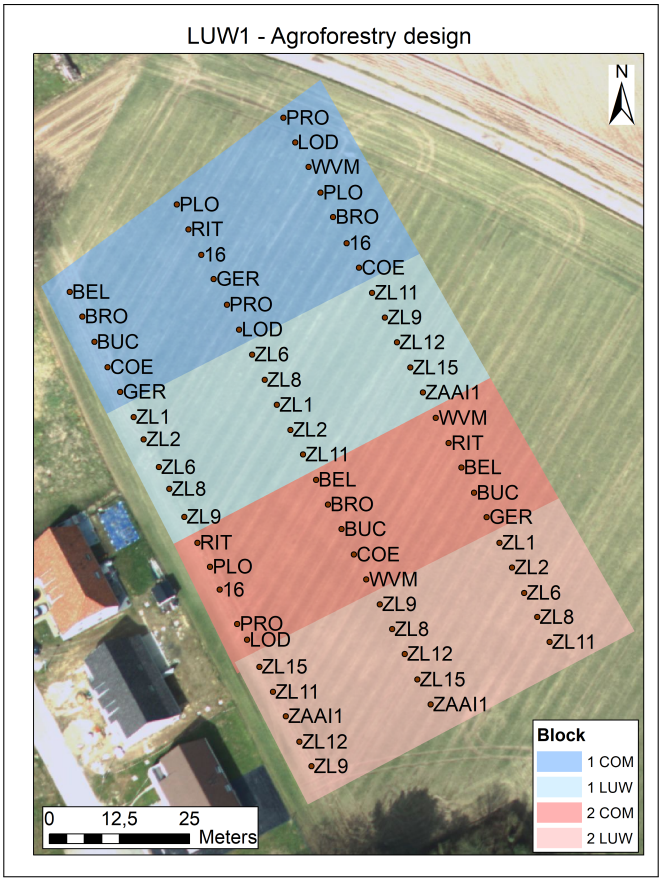


Figure 5 – LUW1 - Agroforestry design.

3.1.2 LUW2 - Herent

The site in Herent, LUW2, is an 0.8 ha cropland plot (figure 6). Only half of this plot is used for the late budding walnut trees experiment. This site is managed by Praktijkpunt Landbouw Vlaams-Brabant research institute.

The soil map (figure 6) shows a gradient in soil conditions. The north of the plot is an Aca. The soil is therefore a loam with moderate drainage and B textural. In the south the soil is a Aba so the only difference is that the drainage is considered as "good". With regard to the topography, the field has no significant slopes.

The design of the experimental plot of Herent site is presented in figure 7. The trial is composed of three rows of 12 walnuts trees. Rows are 15 m apart and trees from the same line are 5 m apart. As for ILVO plot, trees are separated in blocks to separate late budding varieties from commercial ones and in such a way that each variety comes back only once in each row (figure 7).

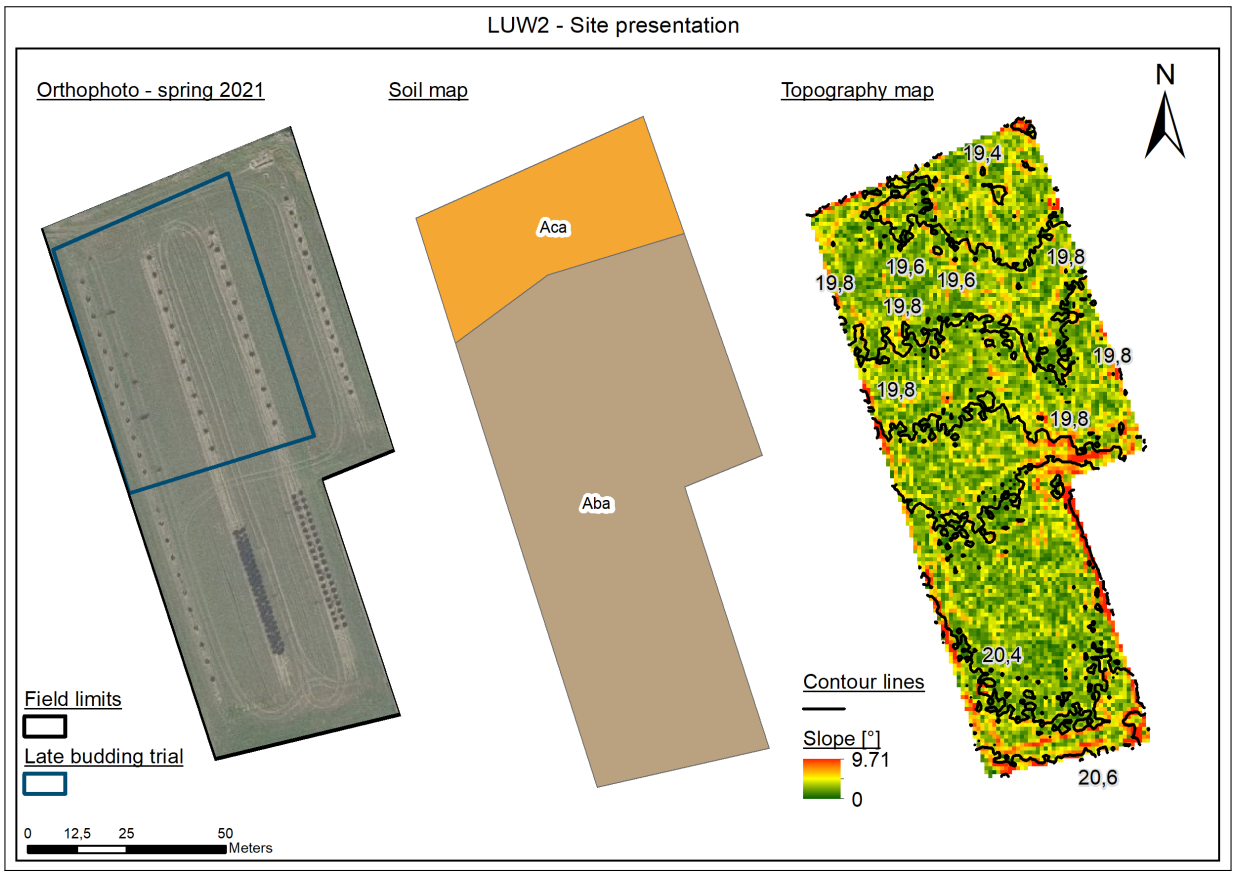


Figure 6 – LUW2 - Orthophoto, soil map and topography (1 m x 1 m).

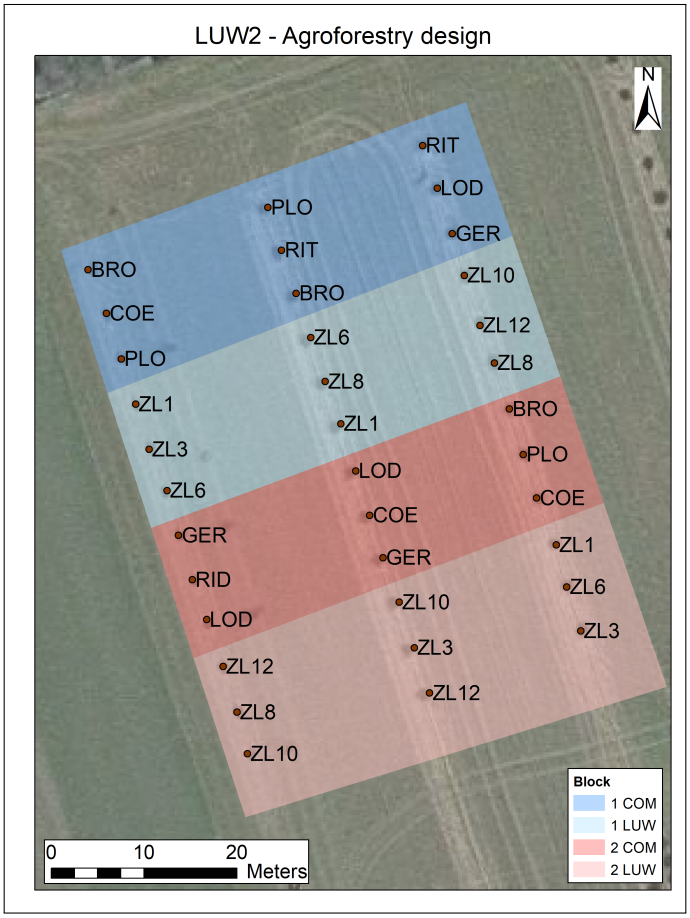


Figure 7 – LUW2 - Agroforestry design.

3.1.3 LUW3 - Galmaarden

Galmaarden site (figure 8), LUW3, is a 5 ha cropland plot owned and managed by a farmer who can be considered an AF pioneer in Flanders.

The soil map is also presented in figure 8. The majority of the site's extent on the eastern side is considered as an Aca1 while the rest is considered as an Acp. The texture is therefore a loam with a moderate drainage but the difference concerns the profile development. Aca1 soil has a B textural horizon with a A horizon less than 40 cm thick and Acp has no profile development.

Concerning the topography, there are slopes of more than 5 % and the difference between highest and lowest points is around 10 m. This explains why there is no profile development in certain zones. Indeed, Acp area is a zone of colluvial deposits due to the slope.

The design of Galmaarden is presented in figure 9. There are four lines of 4, 18, 18 and 17 trees being part of the experiment. There are therefore 57 walnut trees specifically planted for the experiment even if more trees are planted in total on this site. Only LUW trees are represented in fig 9. Rows are 65 m apart and trees from the same row are 5 m apart.

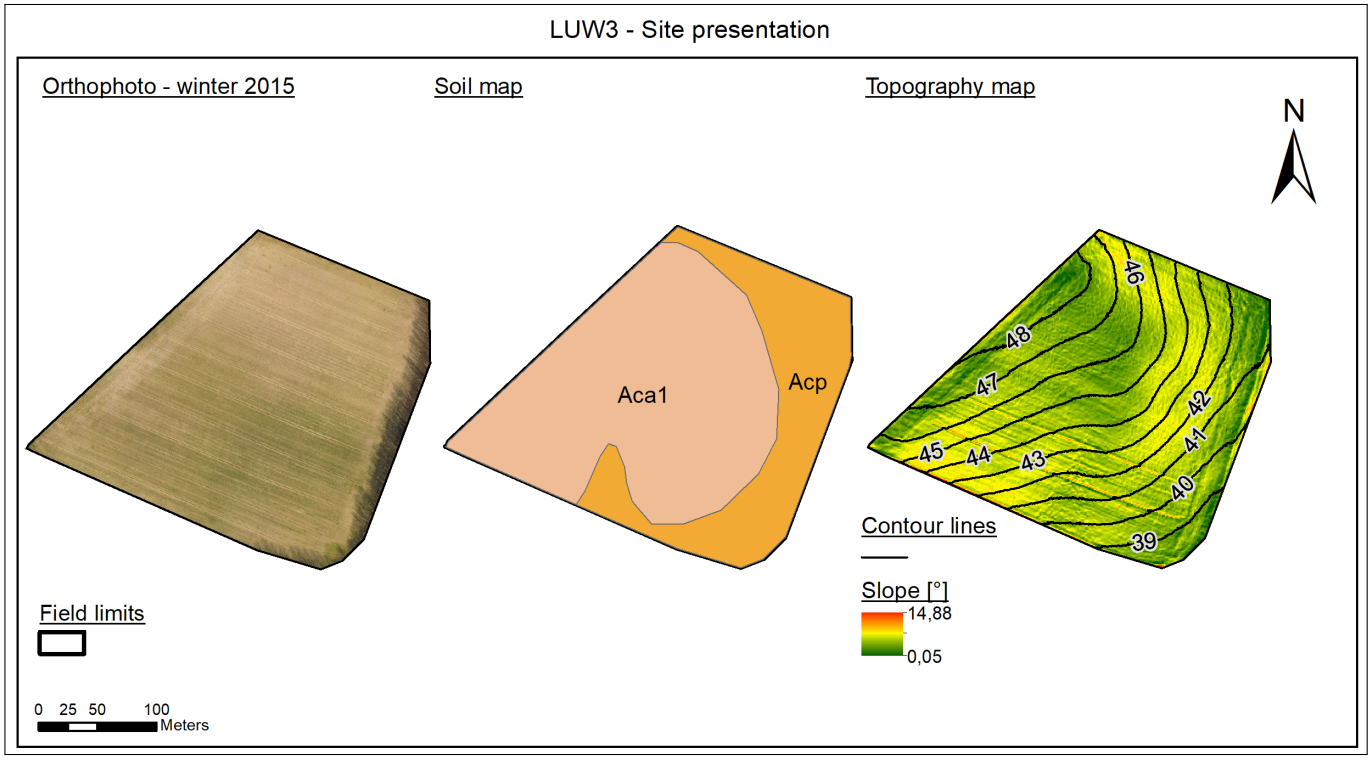


Figure 8 – LUW3 - Orthophoto, soil map and topography (1 m x 1 m).

Figure 9 – LUW3 - Agroforestry design.

3.1.4 LUW4 - Hulste

LUW4 is an 3.1 ha pasture field with cows presented in figure 10. About 0.5 ha concern the late budding walnut trees experiment. This site is owned and managed by a farmer.

There are four different soil types : Pcc, Pdc, Lep, Eep. The first one is a moderately dry light sandy loam soil with strongly mottled and crumbled B_t horizon. The second one is moderately wet light sandy loam soil with heavily mottled and crumbled B_t horizon. Lep means sandy loam with a rather poor drainage, strongly gleyed and with no profile development. Finaly Eep is light clay with a rather poor drainage, strongly gleyed and with no profile development. The experimental trial zone only deals with Pcc, Pdc and Lep. As for Herent and ILVO sites, the ground of Hulste is rather flat.

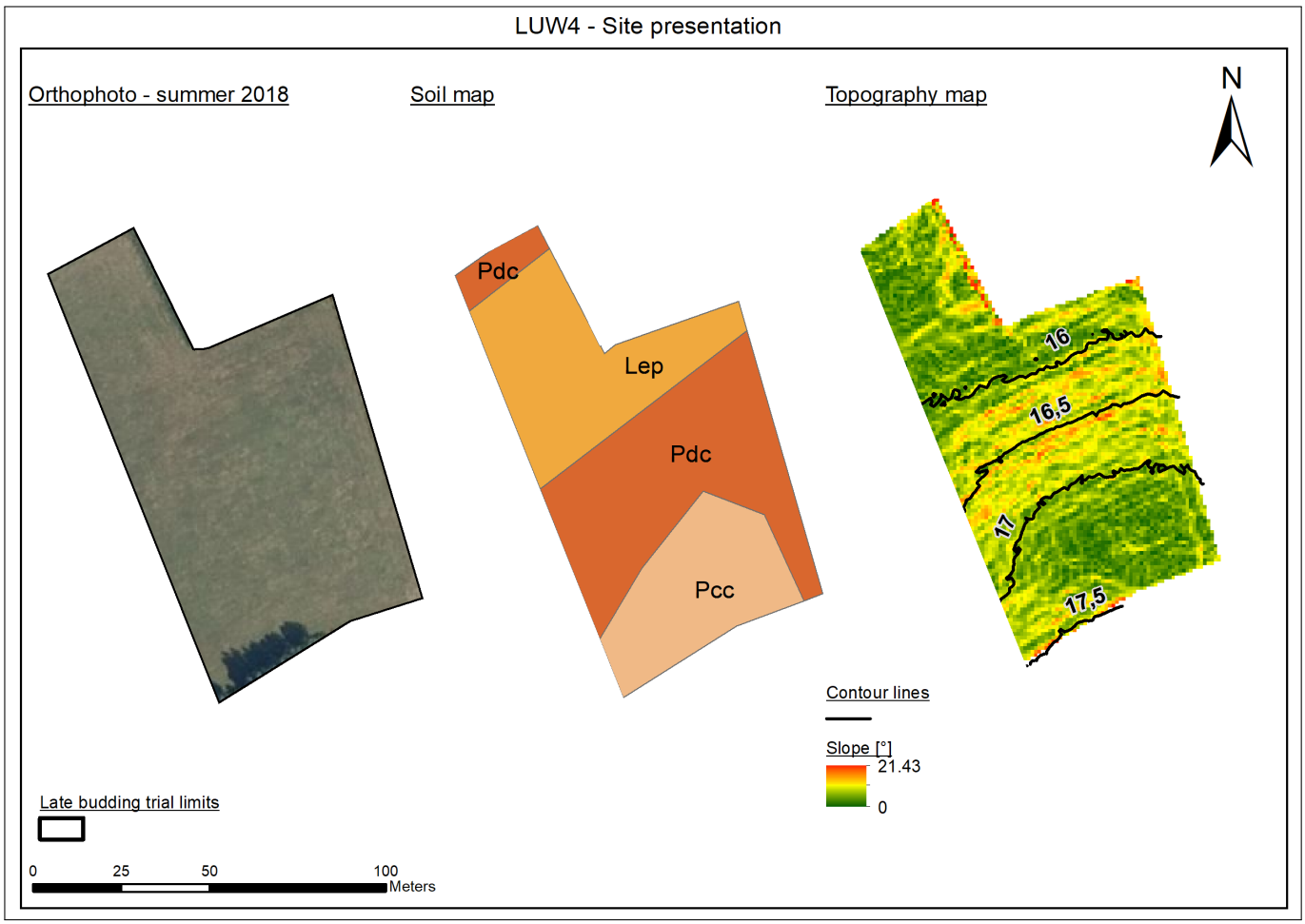


Figure 10 – LUW4 - Orthophoto, soil map and topography (1 m x 1 m).

The design of the late-budding trial part of Hulste is presented in figure 11. The trial is composed of three lines of 7, 7 and 8 walnuts trees. Rows are 14 m apart and trees from the same row are 11 m apart.

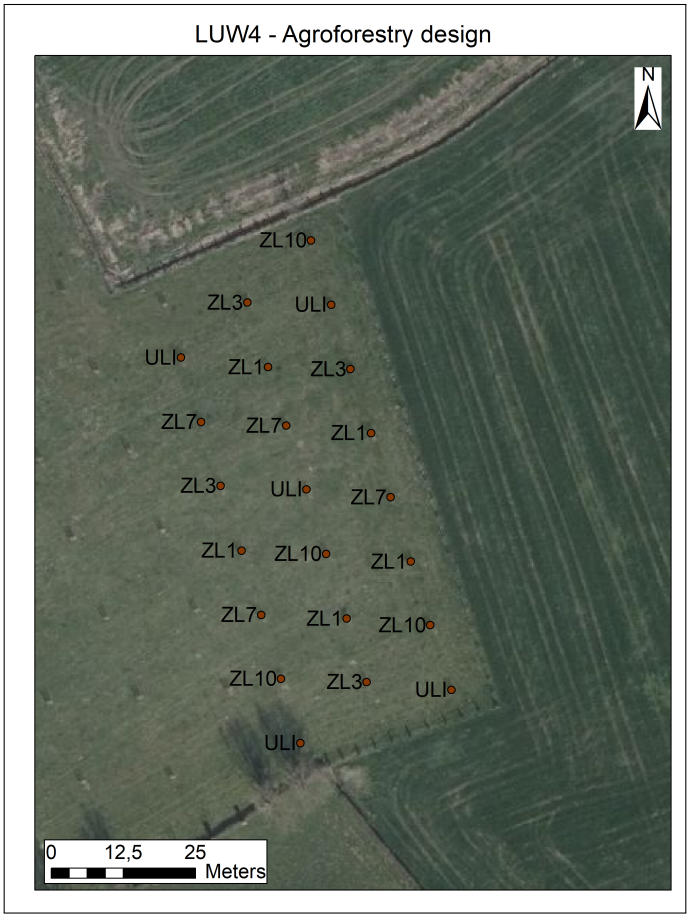


Figure 11 – LUW4 - Agroforestry design.

3.2 Walnut tree varieties

All trees planted in the experimental sites for the project were bought as two year old trees from the tree nursery "De Veredelde Walnoot" (Ulicoten, Netherlands). These young trees are produced using the technique of grafting. In other words, a piece of wood from the parent tree of the respective variety is grafted on top of a rootstock from common walnut (Juglans regia L.) seedlings.

For the commercial varieties, the parent trees used for the propagation originate from other Dutch tree nurseries (Van 't Westeind, Achtplagennuts, De Smallekamp and 't Herenland).

The parent trees for the newly tested varieties (ZL code) are individual trees, spread over the South of the Netherlands and Flanders, which were noticed to have the characteristic of budding very late in the season (i.e. only starting from June or later).

Table 3 shows the code used for commercial walnut varieties. Concerning the full name of the ZL trees, it is confidential at this stage of the project. More information about the full names should be obtained by contacting Bert Reubens.

Table 3 – Walnut tree variety code.

Commercial varieties	
Code	**Name**
BRO	Broadview
BUC	Buccaneer
COE	Coenen
PLO	Plovdivski
RIT	Rita
16	Nr 16
GER	Germisara
PRO	Proslavski
LOD	Lange van Lod
BEL	Bella Maria
WVM	Wunder von Monrepos
MAR	Marcella

3.3 Soil characterisation

3.3.1 Soil mapping with electromagnetic induction

In order to characterise the soil heterogeneity, a mapping and sampling survey has been conducted in the four AF sites on March 2 and 3. Additional sampling has been conducted November 9 2021 to complete the data set.

Prior to taking the soil samples, an EMI scan has been carried out by a team from Ghent University leaded by Prof. Philippe De Smedt (figure 12). The device used for the scanning was a Dualem 21HS, the sampling rate was 10 Hz and the line spacing was 2 m. The drift correction and general data processing flow are based on Delefortrie et al. (2014) and Hanssens et al. (2020) papers. Finally, the method to interpolate the results was the natural neighbour method.

Figure 12 – EMI soil scanning on LUW1.

The EMI sensor used consists of one transmitter coil and six receiver coils located at 0.5, 0.6, 1, 1.1, 2 and 2.1 m distance from the transmitter. There were two different orientations of the coils : horizontal coplanar (HCP) dipole and perpendicular (PRP) dipole modes. The coils located at 0.5, 1 and 2 m from the transmitter were in HCP orientation (HCPH, HCP1 and HCP2) while the other ones were in PRP orientation (PRPH, PRP1, PRP2).

The depth of investigation (DOI), which is the depth below which sensors are insensitive to the value of soil properties (Oldenburg and Li, 1999) is determined by the distance between the transmitter coil and the receiver one. This depth is conventionally defined as the depth above which 70 % of the cumulative response is obtained from the soil volume (Taylor, R., 2005). It is also influenced by the orientation of transmitter-receiver pair. The formula used to calculate the cumulative EC_a response for the the HCP dipole mode (C_{HCP}) and for the PRP dipole mode (C_{PRP}) are taken from Wait (1982) and are presented in equation 8 and 9 :

$$C_{\mathrm{HCP},s}(z) = \frac{1}{\left(4 \cdot \frac{z^2}{s^2} + 1\right)^{\frac{1}{2}}} \tag{8}$$

$$C_{\mathrm{PRP},s}(z) = 1 - \frac{2\frac{z}{s}}{\left(4 \cdot \frac{z^2}{s^2} + 1\right)^{\frac{1}{2}}} \tag{9}$$

With :

C_{HCP} : The cumulative EC_a response from a soil volume above depth z for the HCP orientation (in % of the measured signal, relative to 1)
C_{PRP} : The cumulative EC_a response from a soil volume above depth z for the PRP orientation (in % of the measured signal, relative to 1)
z : Depth [m]
s : separation between transmitter and receiver coils [m]

The DOI are therefore approximated as 0.8 m, 1.6 m and 3.2 m for HCPH, HCP1 and HCP2 respectively and 0.3 m, 0.55 m and 1 m for PRPH, PRP1 and PRP2. Finally the instrument elevation above the ground (ca. 0.16 m) must be subtracted . The real DOI are then approximated as 0.65 m, 1.45 m and 3 m for HCPH, HCP1 and HCP2 and as 0.15 m, 0.4 m and 0.85 m for PRPH, PRP1 and PRP2.

In addition to the DOI, the relative response of the EMI signal as a function of depth, $R(z)$, can also be calculated (equations 10 and 11) (Wait, 1982).

$$R_{\mathrm{HCP},s}(z) = \frac{4\frac{z}{s}}{s\left(4\left(\frac{z}{s}\right)^2 + 1\right)^{\frac{3}{2}}} \tag{10}$$

$$R_{\mathrm{PRP},s}(z) = \frac{2}{s\left(4\left(\frac{z}{s}\right)^2 + 1\right)^{\frac{3}{2}}} \tag{11}$$

By plotting these equations (figure 13), the differences of behaviour can be observed between the coil orientations. The PRP orientation is most sensitive to the conductivity of the surface and are less and less sensitive when the depth increases. For the HCP orientation, the sensitivity firstly increases until it reaches a depth of maximum sensitivity and then decreases. After the subtraction of the instrument elevation above the ground, the depth of maximum sensitivity is equal to 2 cm for HCPH, 20 cm for HCP1 and 55 cm for HCP2. Despite the concept of a DOI and of the depth of maximum sensitivity, it is important to remind that there is still an influence of deeper layer in the EMI response.

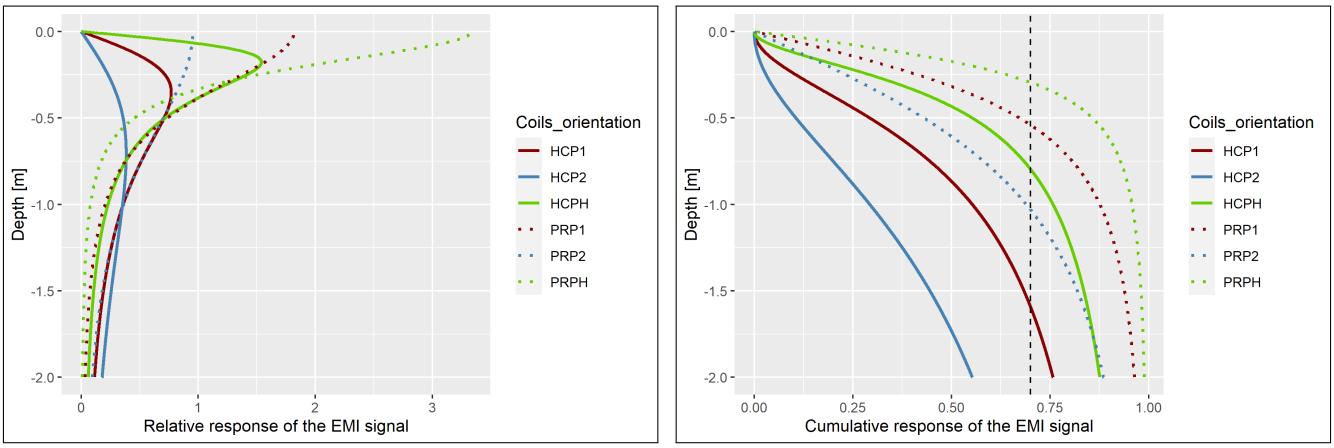


Figure 13 – Relative (a) and cumulative (b) depth response curves of the EC_a for the six EMI coil configurations and with indication of the 70 % cumulative response.

3.3.2 Sampling

In order to determine the position of the soil samples, the conditioned Latin hypercube sampling (cLHS) technique (Minasny and McBratney, 2006) was used. It is a "sampling strategy with prior information represented as exhaustive ancillary data" (Minasny and McBratney, 2006), in this case the ancillary data are the EMI data. This is a random stratified sampling approach with the aim of fully sampling the recorded conductivity variation. In this way, six locations per site were determined for the purpose of collecting soil samples (figure 14, samples A-F). The objective of the soil analysis is to highlight soil properties that explain EC_a. In this way, the homogeneous soil areas identified by clustering can be linked to specific soil characteristics.

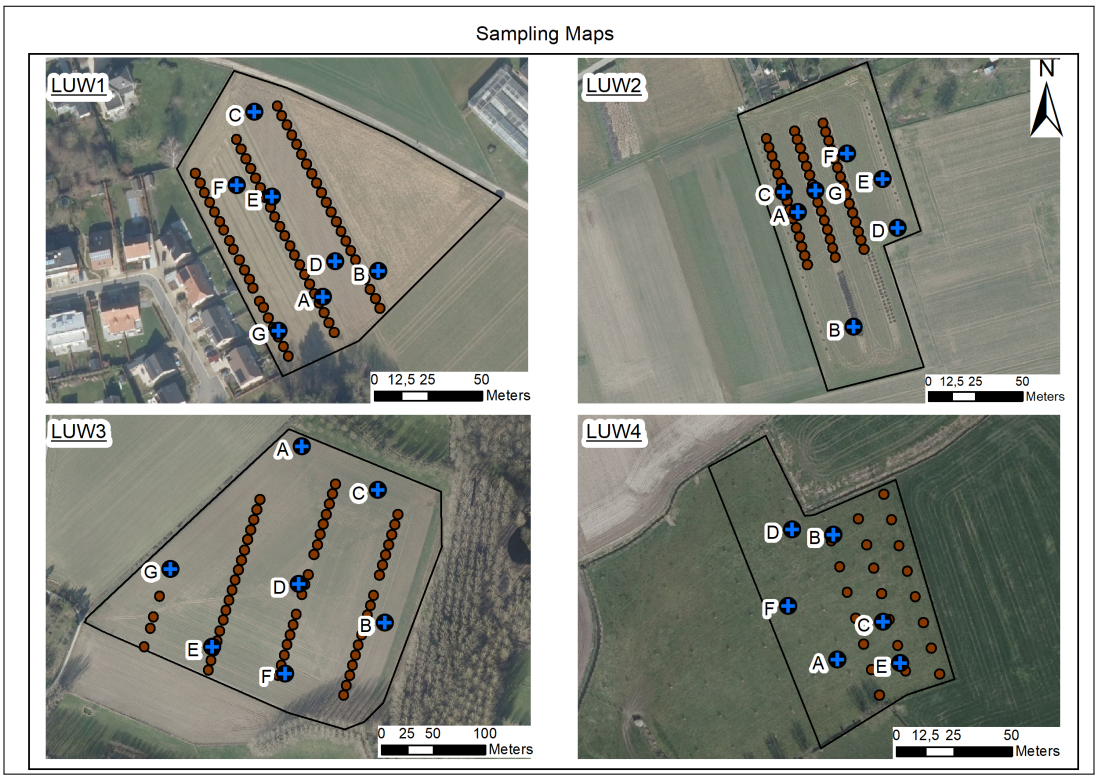


Figure 14 – Sample locations in all sites.

Three types of samples have been taken at two depths for each of the 24 locations. The first type consisted of undisturbed samples taken at 20 and 40 cm in order to measure the soil moisture and the bulk density. They have been taken with Kopecky's rings. The second and third types have been collected with an Edelman Auger. Samples were taken from 0 to 30 cm and between 30 and 60 cm. For each depth, one sample was intended for the laboratory of La Hulpe to analyse the texture and pH H_2O and the second was for the laboratory of ILVO to analyse phosphorus, potassium, calcium, magnesium, manganese, iron and sodium with ammonium acetate extraction and soil organic carbon (SOC), cation exchange capacity (CEC), total nitrogen and pH KCl. The method used by the laboratories are resumed in table 4.

During the month of November, the determination of the clusters based on the EMI data showed that it was necessary to take new samples at specific locations (figure 14, sample G) in LUW1, LUW2 and LUW3. The reason is that there were no samples in some clusters.

Table 4 – Method for the soil analyses.

Analysis	**Laboratory**	**Method**
Texture	La Hulpe	Sedimentation and sieving, derived from the NF X 31-107 standard.
pH H2O	La Hulpe	ISO 10390
Phosphorus	La Hulpe	Determination by injection flow colorimetry Own method with extraction from Lakanen and Erviö (1971)
Potassium Magnésium Calcium	La Hulpe	Determination by flame atomic absorption spectrometry Own method with extraction from Lakanen and Erviö (1971)
SOC N total	ILVO	Dry combustion at 1100 °C using a Skalar C/N-Analyzer Primacs SNC100-IC ISO 10694 for SOC and ISO 13878 for N
Iron Phosphorus Potassium Magnésium Calcium Manganese Aluminium Sodium	ILVO	Soil extraction with 0.1 M ammonium lactate + 0.4 M acetic acid at pH 3.75 Measurement by inductively coupled plasma optical emission spectroscopy (Egnér et al., 1960)
CEC	ILVO	Determined by ammonium acetate at pH 7.0 and KCl
pH KCl	ILVO	ISO 10390

Soil moisture analysis has been carried out on March 2 and 3.

Every undisturbed sample has been removed from the Kopecky' ring in a weighted aluminium container and then has been weighted in fresh conditions. They were then dried at 105 °C in the oven for 24 hours. Finally, all samples have been weighted again after drying.

The formulas used for gravimetric and volumetric soil moisture and for bulk density are given in equation 12, 13 and 14.

$$\mu = \frac{m_l}{m_s} = \frac{m_f - m_d}{m_d} = \frac{m_{tf} - m_{td}}{m_{td} - m_c} \tag{12}$$

$$\theta = \frac{V_l}{V_t} = \frac{m_f - m_d}{V_s} \tag{13}$$

$$BD = \frac{m_s}{V_s} = \frac{m_{td} - m_c}{V_s} \tag{14}$$

With :

μ : gravimetric moisture content [-]
θ : volumetric moisture content [-]
BD : bulk density [g/cm^3]
m_l : liquid mass [g]
m_s : solid mass [g]
m_f : fresh sample mass [g]
m_d : dry sample mass [g]
m_{tf} : total fresh sample mass (weight of the container included) [g]
m_{td} : total dry sample mass (weight of the container included) [g]
m_c : aluminium container mass [g]
V_l : liquid volume [cm^3]
V_t : total volume = 100 cm^3

With the purpose of comparing the results of the soil sampling to the results of the EMI scan, a mean value of a circle with a diameter of 1 m have been calculated around each sample locations for each EMI scan raster.

3.4 Soil moisture dynamics in ILVO site

3.4.1 Measurements

Three soil moisture content probes (Sentek, 2021) have been installed on the ILVO site for the purposes of analysing moisture dynamics in three locations into the walnut trees rows. It will turn out later that each probe corresponds to a specific cluster.

Decision for the probe locations has been based on EMI data and topography of the site and can be seen on figure 15. The first decision was to install one probe by row because clear EC_a tendencies can be observed between each row. Then for the location within the row, they follow the natural topography of the site.

These probes measure water content by volume based on variations in the soil dielectric constant over time. The three probes give soil moisture content in percentage of the volume (cm^3_w / cm^3_{soil}) at three depths (25, 55 and 85 cm) and every 30 minutes. The probes have not been specifically calibrated for the ILVO site, the equation for transforming the measurement into water content having remained the default one (equation 15). This default calibration equation is sufficient to show soil moisture trends and can be therefore used to study relative soil water changes (Sentek, 2011). Measurements started on March 25 at 3 :30 pm and were made on a 30-minute time step. Data were frequently collected until 09/11/2021.

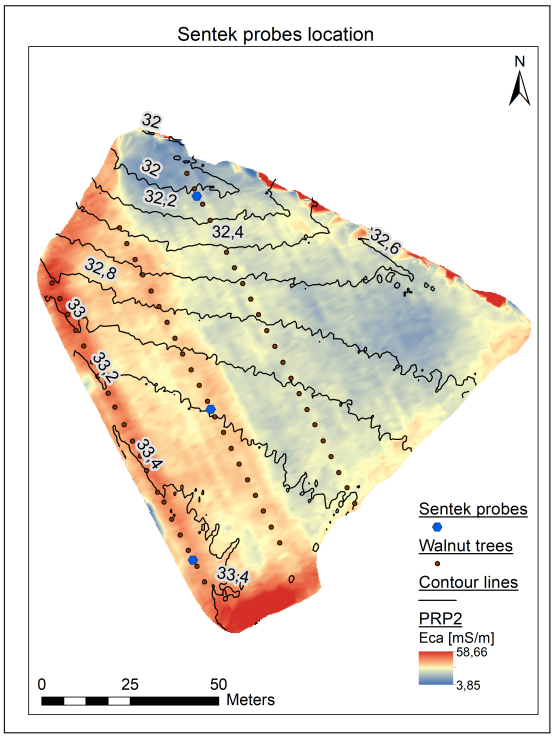


Figure 15 – Sentek probes locations.

$$y = 0.1957x^{0.4040} + 0.02852 \tag{15}$$

With :

y : scaled frequency (reading from the access tube) [-].

x : volumetric moisture content [mm].

Some technical problems occurred during the experiment. The probe in the third row has measurements only for 25 and 85 cm during the all period of measurement. Moreover the probe in the middle row has sometimes a lack of data. Indeed between the 22-04 and the 28-04, between the 11-08 and 02-09 and after the 27-09, data are missing.

It is important to note that absolute values should be handled with care, since different soil horizons and different locations might have distinct calibration relationships.

3.4.2 Evapotranspiration calculation

In order to analyse the soil moisture content data, and to validate them, reference evapotranspiration (ET_0) has been calculated on a daily time step. ET_0 is defined as the evaporation of a hypothetical grass reference crop without water stress (Allen et al., 1998). The software *ET_0 Calculator* has been used to calculate the evapotranspiration via the Penman-Monteith formula (Raes and Munoz, 2009).

Meteorological data used to calculate ET_0 for ILVO site have been provided by the Royal Meteorological Institute. It is gridded observational data covering Belgium at a spatial resolution of 5 km. Variables are temperature minimum and maximum [° C], precipitation quantity [mm], relative humidity [%], wind speed at 2 m [m s^{-1}] and global solar radiation [MJ m^{-2} day $^{-1}$] for a daily time-step.

3.5 Clustering

With the aim to determine a limited number of management zones on each experimental sites, a machine learning algorithm has been used. We used the spatial fuzzy c-means (SFCM) clustering method (Cai et al., 2007) implemented by Gelb and Apparicio (2021) in the R package *geocmeans*.

The raster used in the algorithm are HCPHQP, HCP1QP, HCP2QP, PRPHQP, PRP1QP and PRP2QP. However, the rasters were not used in their entirety and the raw data were transformed. The goal was to avoid outliers. Concerning LUW1 and LUW4, rasters were clipped to focus only on the experimental part of the site. For LUW1, the edges of the rasters were very noisy. Moreover, there was a different soil cover in the eastern part of the field, which biased the results of the scan. Concerning LUW4, rasters were also clipped around the trees. Indeed, during the scan of the soil, there was a metal fence around the trees to protect them. This distorted the scan results. Finally, outliers were treated by reducing them to $\mu - 3\sigma$ or $\mu + 3\sigma$ with μ the mean of the EC_a distribution and σ its standart deviation. The decision to not delete the outliers is to keep the same pixels for every rasters, which is necessary for the SFCM in *RStudio*.

To optimise the choice of the algorithm parameters values, three indexes have been taken into account. The first one is the silhouette value (Rousseeuw, 1987). With this method, each data point has a silhouette that is based on the dissimilarities (how much the data point of a cluster are separated from the data points of other clusters) and the similarities (how much the data point is similar to those of its own cluster). This value is between -1 and 1. A positive value depicts a situation where the data point is correctly classified and a negative value means that it is not in the correct cluster (von Hebel et al., 2021). The higher the value, the greater is the cohesion separation. A value of 0 indicates that the cluster contains only one object (Rousseeuw, 1987). The second index used is the explained inertia of the classification. Formulas to calculate this index are given in equation 16, 17, 18 and 19 (Institut de Mathématiques de Bordeaux, 2015). This criterion varies between 0 and 100 %. A value of 0 means that there is only one cluster and a value of 100 is for a partition in n clusters, with n the number of individuals (Institut de Mathématiques de Bordeaux, 2015). This index will be mainly used to set the number of groups by looking at the gains of explained inertia between the different classifications (Gelb and Apparicio, 2021). Finally, the last used index is Xie-Beni (XB) index (Xie and Beni, 1991). This index is the ratio of compactness and separation of a fuzzy

classification (Gelb and Apparicio, 2021). A low value is desirable for this indicator. Formulas for XB are given in equation 20, 21 and 22 and are adapted from Singh et al. (2017).

$$EI = \left(1 - \frac{W}{T}\right) \times 100 \tag{16}$$

$$W = \sum_{k=1}^{K} I\left(C_k\right) \tag{17}$$

$$I\left(C_k\right) = \sum_{i \in C_k} \frac{1}{n} d^2\left(x_i, g_k\right) \tag{18}$$

$$T = \sum_{i=1}^{n} \frac{1}{n} d^2\left(x_i, g\right) \tag{19}$$

With :

EI : explained inertia [%].

W : intra-class inertia.

T : total inertia.

g : gravity center of the cloud.

$I\left(C_k\right)$: inertia of one class.

K : number of classes of the classification.

x : vector associated with one individual.

n : number of individuals.

d : euclidian distance.

$$J_m = \frac{1}{N} \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij}^m d^2\left(X_j, c_i\right) \tag{20}$$

$$d_{\min} = \min_{i,j} \left[d^2\left(c_i, c_j\right)\right] \tag{21}$$

$$XB = \frac{J_2}{d_{\min}} \tag{22}$$

With :

J_m : the compactness of a fuzzy cluster.

N : the number of individuals.

u_{ij} : the fuzzy membership value of an individual.

m : the fuzziness degree parameter.

$d^2(X_j, c_i)$: the distance between the cluster center and the data point .

$d_{\min}$: The minimum distance between the cluster centroids .

XB : Xie-Beni index.

To fix the number of class (k) and the fuzziness degree parameter (m) of the clustering, 120 SFCM has been computed for each site. Indeed, the tree parameters mentioned above has been calculated for each value of k from 2 to 5 and for each value of m from 1.1 to 4 with a step of 0.1 . Analysis of the index graphs as a function of m and k led to select one m and one k per site.

The next step was to decide on the weight of the spatial data in the SFCM. Therefore, the α parameter and the spatial matrix had to be set. SFCM were computed for each α value between 0.5 and 2 with a step of 0.5. Regarding the spatial matrix, it is an n x n matrix composed of 1 with n an odd number. Five values of n were tested : 21, 31, 41, 51 and 61. Unlike k and m, the choice of α and the spatial matrix has not been made on the basis of the graphs. Indeed, the purpose of these parameters is to avoid noise in the classification by preventing points from being classified differently from all neighbouring points and the best way to investigate this noise is to look directly at the resulting SFCM rasters. Parameters that gave the best rasters have been chosen. Another solution would have been to calculate the spatial consistency index (Gelb and Apparicio, 2021). However, this index is more suitable for shapefile data and the computation time is extremely long for raster data with tens of thousands of data rows.

For the sake of synthesis, the choice of the different parameters of SFCM algorithm is explained only for LUW1. A focus is first done on the appropriate number of cluster. Then, the choice of the fuzziness degree parameter, m, is described. Finally the spatial parameters are chosen.

First, logically the explained inertia (figure 16) increases with the number of groups. We can then see that the gain of inertia is limited after 3 or 4 groups. Concerning the silhouette value (figure 16), the fewer groups, the higher it is. There is no large differences between 2 and 3 groups while the SFCMs with 4 or 5 groups have silhouette value significantly lower. Finally, SFCMs with 2 and 3 groups have the best XB index (figure 18). As the explained inertia is much better for 3 groups than for 2, the number of groups chosen is 3.

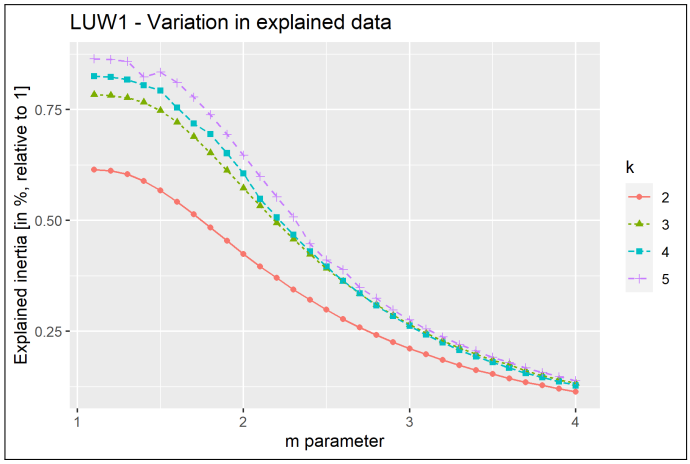


Figure 16 – LUW1 - Explained inertia.

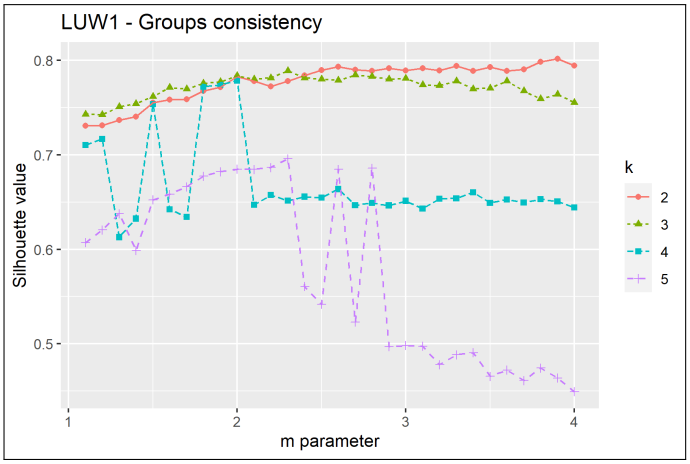


Figure 17 – LUW1 - Silhouette.

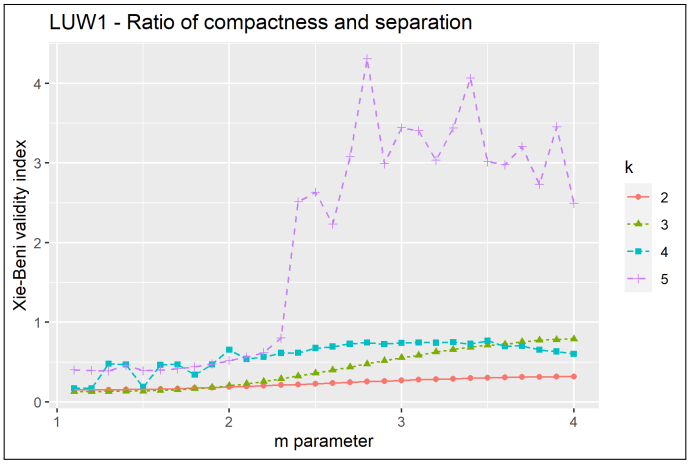


Figure 18 – LUW1 - Xie-Beni index.

The parameter m has been set at 2 for all the sites. It is in fact a usual value of this parameter, used in a lot of studies, and it can be used without particular constraints (Pal and Bezdek, 1995). Then, we see that, when m $\in [1.5, 2.5]$, which is often considered as the optimal range (Pal and Bezdek, 1995), the silhouette value increases with m increasing while on the contrary the explained inertia decreases. Concerning the XB index, it is quite stable in this range of m for LUW1 (figure 18) but also for the other sites. It is therefore difficult to make a choice in this range of m. Moreover, this small changes don't really modify the trends in the delimitation of the management areas which is the objective behind the clustering.

With regard to the choice of α and of the spatial matrix, it has been the same for each experimental site, α=2 and the spatial matrix is a square matrix of 61 pixels. Indeed, whatever the value of α and the size of the matrix, trends are always the same for the results. The only difference is the intensity of the noise. These two parameters have therefore been set at rather high values and the results were then sieved to avoid noise. Getting spots with a small number of pixels within an area of another cluster may make sense in brain imaging, for which this algorithm was first used, but for agricultural purposes it is not relevant.

Finally, the same methodology was applied to normalized difference vegetation index (NDVI) data in addition of PRPH and PRP1 data. The objective is to compare the results of this second clustering with the first results. Since the objective is to determine homogeneous areas in order to interpret the performance of trees in different fields, it makes sense to include a dataset related to the performance of plants on these same plots. The formula of NDVI is adapted from Tucker (1979) (equation 23).

$$\text{NDVI} = \frac{R_{\text{NIR}} - R_{\text{RED}}}{R_{\text{NIR}} + R_{\text{RED}}} * 100 + 100 \tag{23}$$

With :

R_{NIR} : the reflections in the near infrared.

R_{RED} : the reflections in the red.

The data used are Sentinel-2 images obtained on *Google Earth Engine* and dated 24-06-20. The resolution of the images is 8 m x 8 m. In order to have the same resolution of the EMI rasters, the function "disaggregate" from the R-package *raster* has been used with the option "method=bilinear" in order to locally interpolate the values.

3.6 Statistical analysis of the budding time and health of walnut trees

The growth, the budding time and the health of the leaves of the walnut trees have been measured during 2021 by the ILVO. It has been to decide to not analyse the growth in relation to the clusters. The trees were planted early 2021 and this analysis would be too soon to draw conclusions and the growth will still depend too much on the initial state of the tree at planting.

To measure the budding time, a integer score between 0 and 4 was given for each tree once per week between 19-04-21 and 15-08-21. This score is based on the budding stages of the walnut tree with 0 corresponding to a winter bud and 4 corresponding to fully unfolded and fully visible leaves, mostly green in colour (figure 19). Moreover, a number corresponding to the first week with a score different of 0 was given to represent the budding time. It is possible to go from 0 a week to more than 1 the week after and that is the reason why the budding time is considered to be reached when the score is different from 0 instead of being considered to be reached when the score is equal to 1.

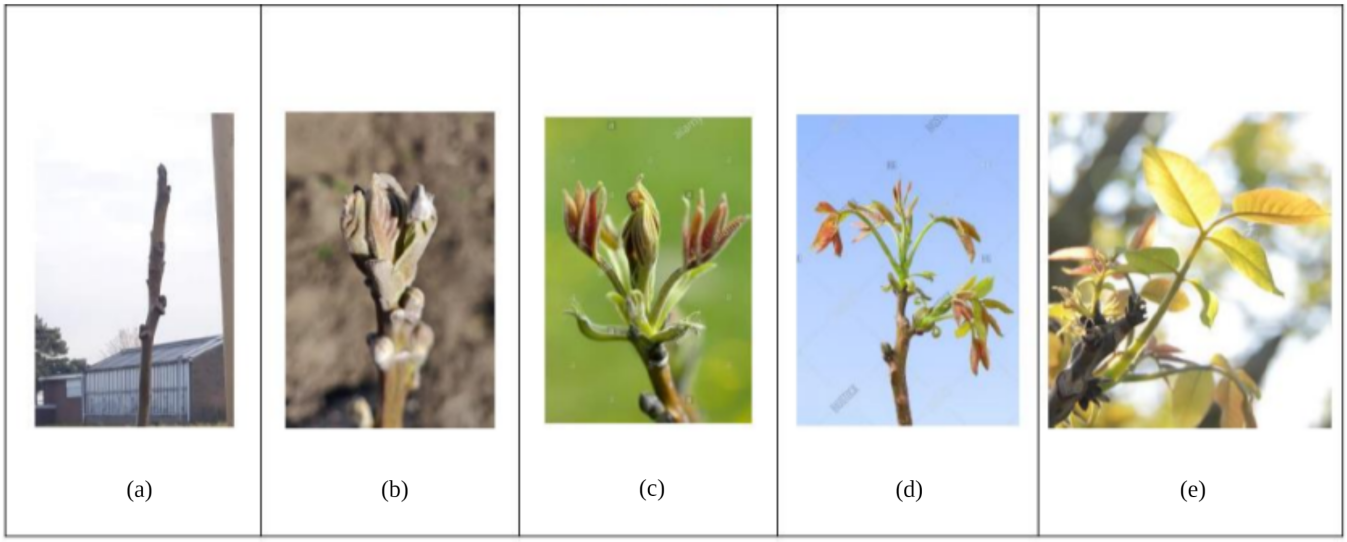


Figure 19 – Budding stages : (a) winter bud ; (b) budburst ; (c) leaflets individualisation ; (d) leaves widely visible but not yet unfolded ; (e) fully unfolded and fully visible leaves mainly green.

Concerning the health of the trees, a integer score between 0 and 4, as for the budding stages, was given to measure the damage of the leaves for each tree. The score of 0 correspond to no damage and a score of 4 correspond to a walnut tree with nearly all its leaves completely damaged. The date of these measures are 23-06-21 and 13-09-21 for LUW1, 29-06-21 and 30-09-21 for LUW2, 28-09-21 for LUW3 and 27-09-21 for LUW4.

The statistical analysis was performed using *RStudio* software (version 4.0.3). Firstly, the normality of each group have been tested using the Shapiro-Wilk normality test from the R-package *rstatix*. Then, the homogeneity of variances between the different group have been tested using Levene's test from the R-package *car*. Then, in order to compare the results as a function of the cluster in which the tree is planted, a one-way ANOVA have been computed for each variable and site with the R-package *rstatix*. Finally, the Tukey's test for post-hoc analysis was used to perform multiple pairwise comparisons between groups (R-package *rstatix*).

4 Results

4.1 Soil mapping and determination of homogeneous zones

4.1.1 Raw EMI maps

Six maps by sites corresponding to the EC_a for all coil geometries are presented in figures 20, 21, 22 and 23.

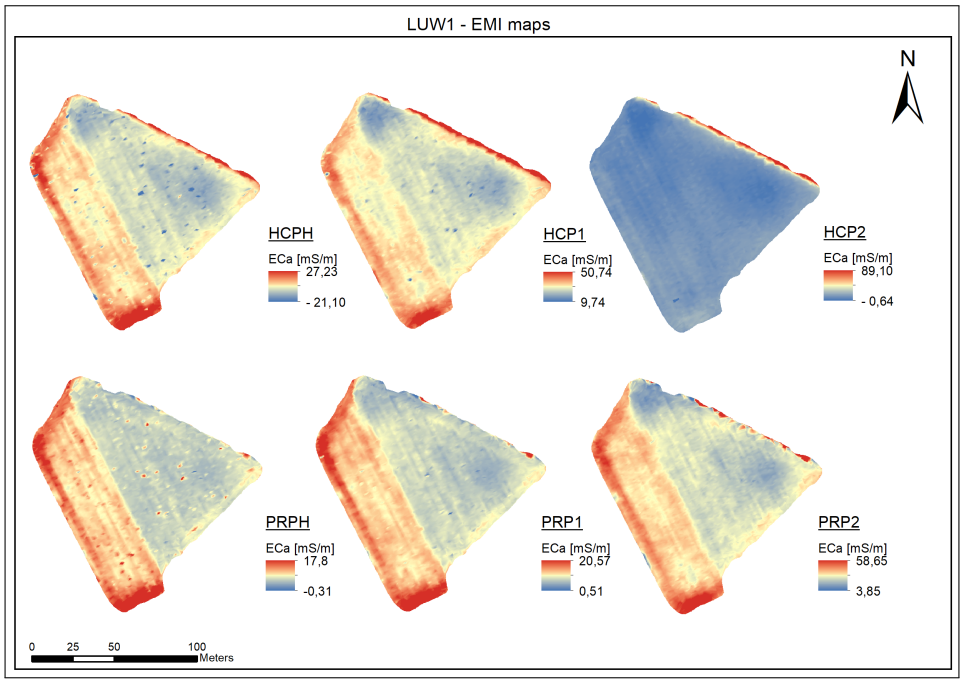


Figure 20 – LUW1 - Maps of the EC_a for each coil configurations (HCPH, HCP1, HCP2, PRPH, PRP1 and PRP2).

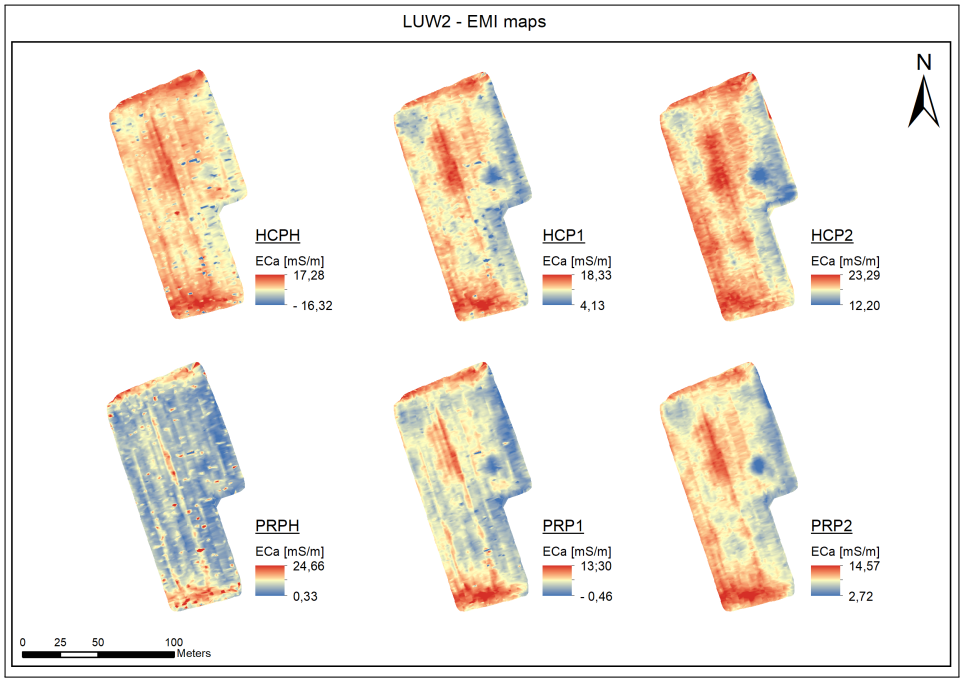


Figure 21 – LUW2 - Maps of the EC_a for each coil configurations (HCPH, HCP1, HCP2, PRPH, PRP1 and PRP2).

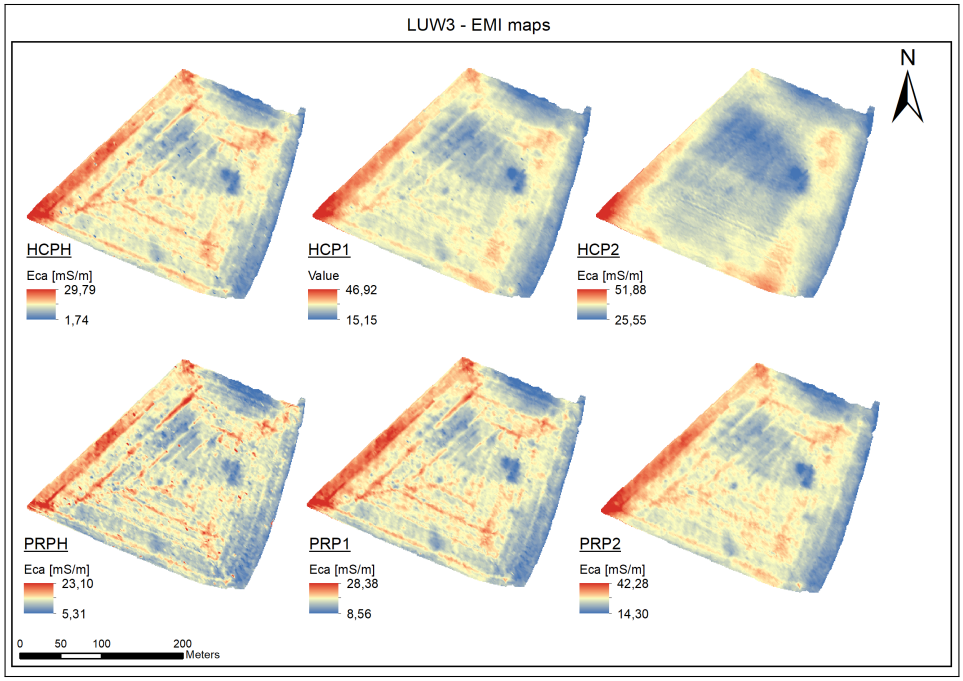


Figure 22 – LUW3 - Maps of the EC_a for each coil configurations (HCPH, HCP1, HCP2, PRPH, PRP1 and PRP2).

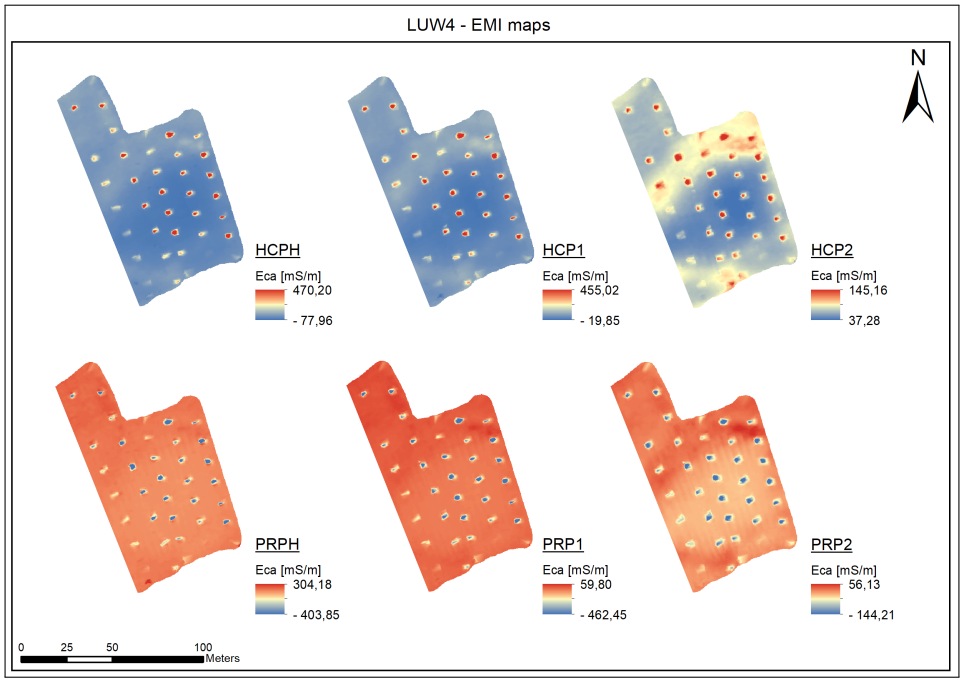


Figure 23 – LUW4 - Maps of the EC_a for each coil configurations (HCPH, HCP1, HCP2, PRPH, PRP1 and PRP2).

4.1.2 Clustered EMI maps

The figure 24 presents the results of the spatial fuzzy c-means algorithm for all the experimental sites. The resulting clusters are considered as homogeneous soil areas.

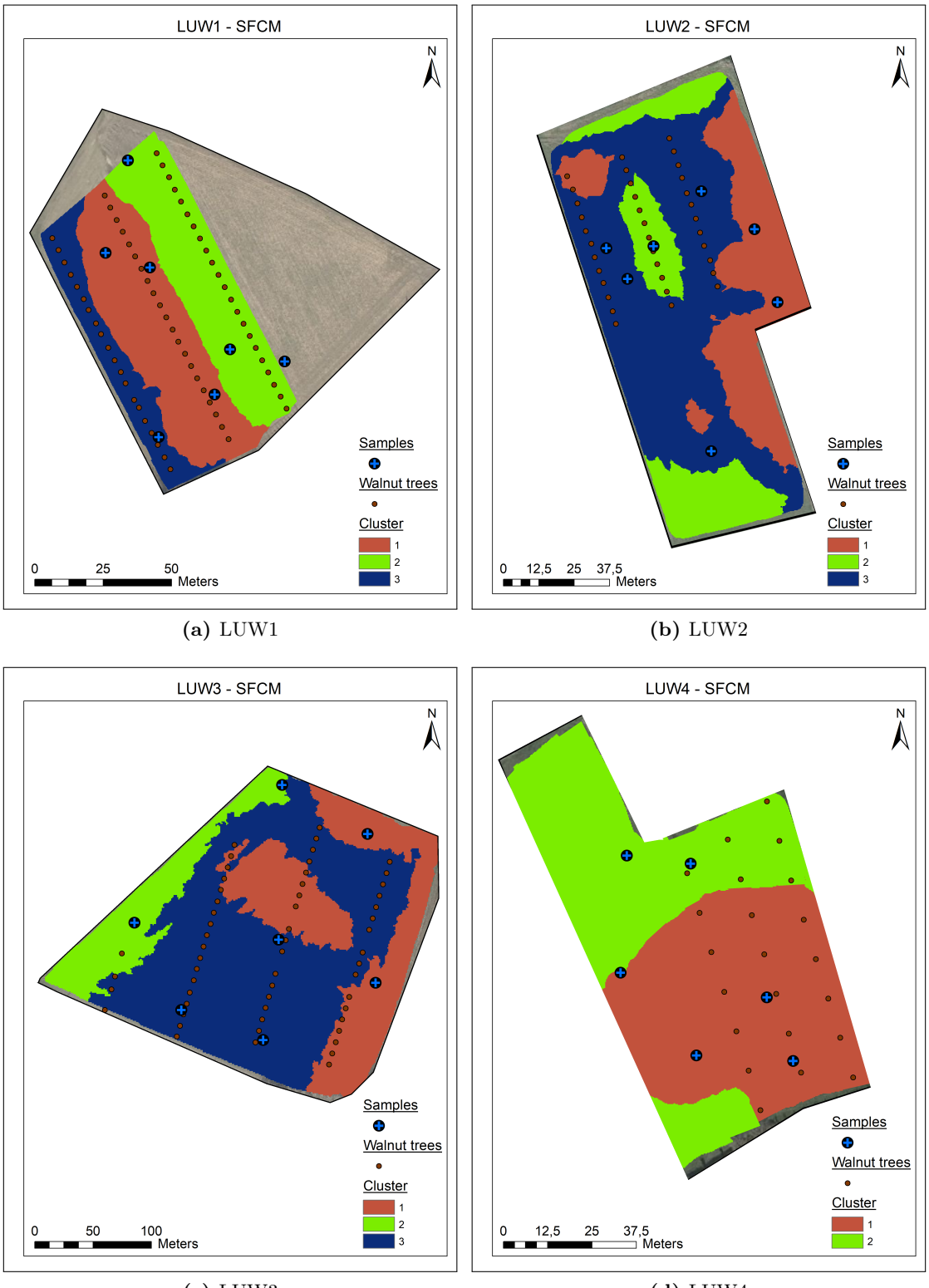
For LUW1, the three identified clusters follow the three walnut trees rows (figure 24a). It should be noted that this was not the intention and that it is simply the result of the clustering of the EMI data. However, this result must be discussed. The clear separation between cluster 1 and 2 is partly due to a different land cover and soil management during the scanning in March. At this moment there were three different land

covers. The eastern part of the field has not been implemented in the SFCM, precisely in order to avoid that the soil cover takes too much importance in the final result and in addition to focus only on the part of the field with the walnut trees. Then the part of the field corresponding the the cluster 2 was tilled while the rest of the field was cultivated.

The figure 24b presents the SFCM results of LUW2. The walnut tree rows are mainly within two of the three clusters. The aim was to do the classification on the whole field instead of just the experimental part of the plot. The first reason is that LUW2 is a small plot and it was interesting to consider all the heterogeneity of the plot and not just the very small area with late-budding walnut trees. Then, SFCMs with a focus on the experimental part gave approximately the same results.

LUW3 SFCM results (figure 24c) highlight four zones (in three clusters) that can be linked with the soil types and the topography of the site (figure 8). First, the cluster 1 zone in the Eastern part of the field matches the colluvial deposits zone (Acp). On the contrary, the second cluster corresponds to the highest area of the field. The third cluster includes approximately the areas of steeper slopes. However the cluster 1 zone in the middle of the plot includes also, between the third and the fourth tree row, a zone with steep slopes.

LUW4 is so the only site with only two clusters (figure 24d). The second cluster, which is mainly in the Northern part of the plot, corresponds to a higher electrical conductivity zone (figure 23). Indeed, in the zone, the soil type is a Lep and therefore the soil has a poor drainage (figure 10). Moreover, the topography shows that this zone is lower than the Southern part (figure 10).



(a) LUW1 **(b)** LUW2

(c) LUW3 **(d)** LUW4

Figure 24 – Clustered EMI maps ((a) LUW1 ; (b) LUW2 ; (c) LUW3 ; (d) LUW4).

4.2 Relationships between EMI signal and soil properties

Scatter plots of EC_a versus soil analysis results were made for the PRP1 and HCPH orientations. The objective is to highlight the soil properties that explain the soil electrical conductivity in order to analyse the clusters produced by the spatial fuzzy c-means classification. The investigation depth of PRP1 orientation (40 cm) approximately corresponds to the 0-30 cm samples. On the other hand, it was not possible to match

the samples taken between 30 and 60 cm to the EC_a of a precise orientation because we are talking about "apparent" soil electrical conductivity and therefore about an average over the entire volume of soil explored. HCPH (DOI = 65 cm) data are therefore compared with a mean value of the soil properties between 0 and 60 cm.

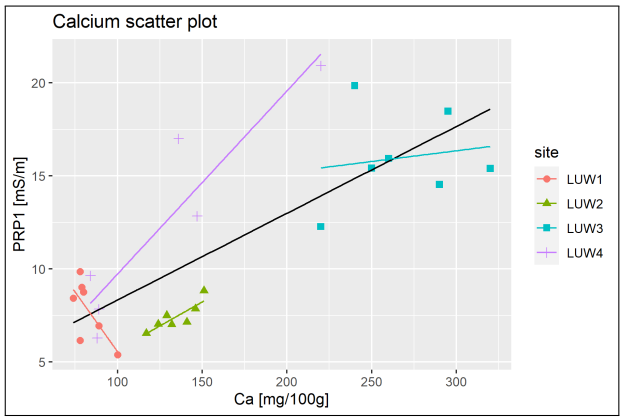
First, correlations has been computed and table 5 gives the soil properties with correlations higher than 0.6 with PRP1 signal and R^2 and p-value of a simple linear regression between the soil property and the EMI signal while table 6 concerns HCPH. Scatter plots of this soil properties are shown in figure 25 for PRP1 and figure 26 for HCPH.

As expected, the texture, via the clay content and the coarse sand content, and the moisture content influence electrical conductivity. CEC is related to texture so it also influence the EC_a. Finally iron, calcium also influences EC_a.

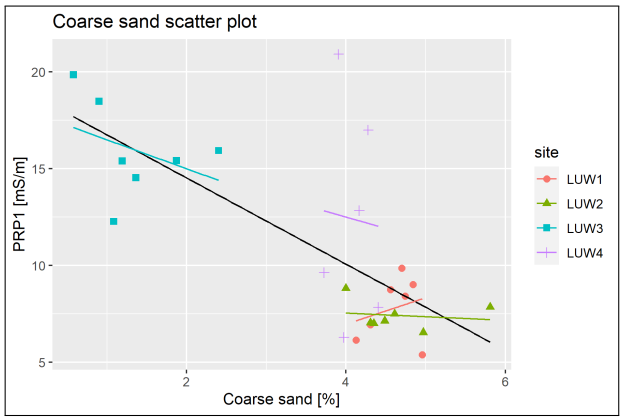
However, LUW4 has a totally different behaviour. EC_a values in this site are higher and slopes in scatter plot are higher than for the other sites. This can be explained by a higher soil iron content, especially in the first 30 cm of the soil. Indeed, figure 27 shows the iron content for the 30 first cm of soil in every site and LUW4 samples values are clearly higher. Figure 28 presents the scatter plots of PRP1 and HCPH versus iron content in LUW4. The iron content is what mainly explains EC_a for this site and that's why LUW4 must be analysed independently from the other sites. Moreover, LUW4 is a pasture while LUW1, LUW2 and LUW3 are cropland which may explain the difference in behaviour of the EMI signal.

Table 5 – Correlations between soil EC_a of PRP1 coil geometry (DOI = 40 cm) and soil properties and R^2 and p-value of the simple linear regression for soil properties with r>0.6.

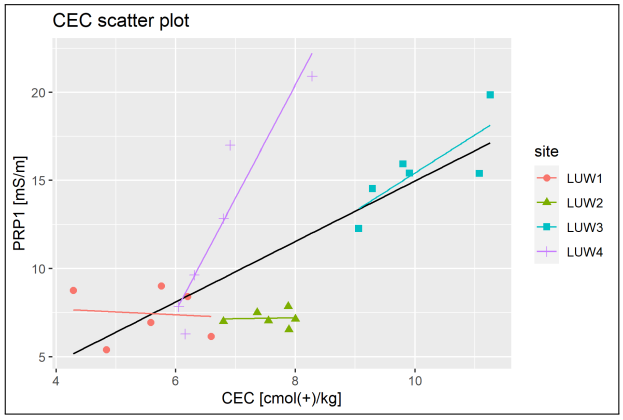
	r	**R^2**	**p-value**
Calcium	0.78	0.60	<0.001
Coarse sand	-0.71	0.50	<0.001
CEC	0.68	0.46	<0.001
Clay	0.63	0.40	<0.001



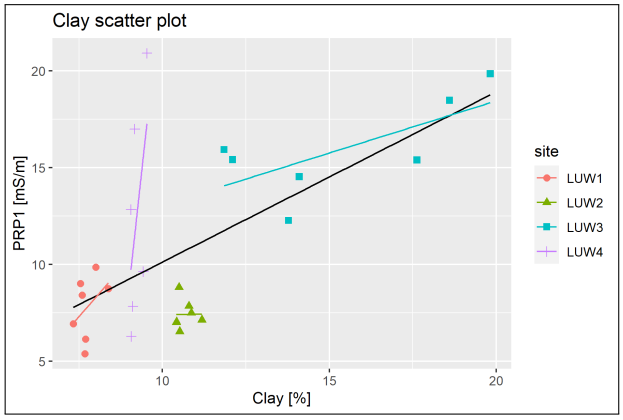
(a) Scatter plot of PRP1 vs calcium



(b) Scatter plot of PRP1 vs coarse sand



(c) Scatter plot of PRP1 vs CEC

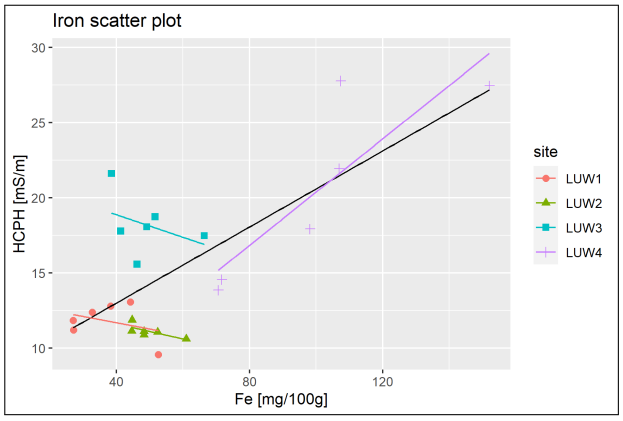


(d) Scatter plot of PRP1 vs clay

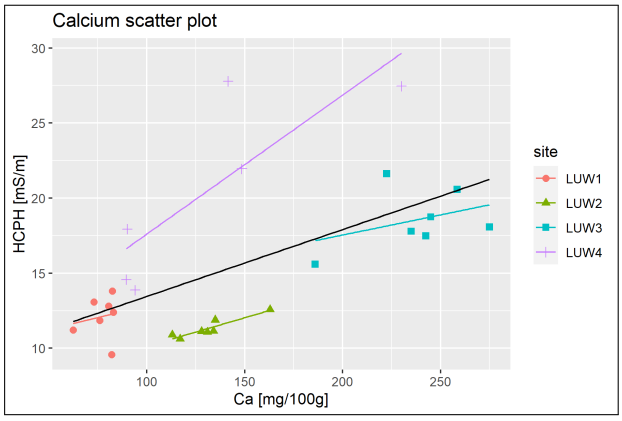
Figure 25 – Scatter plots of PRP1 vs soil properties ((a) Calcium ; (b) Coarse sand ; (c) CEC ; (d) Clay).

Table 6 – Correlations between soil EC_a of HCPH coil geometry (DOI = 65 cm) and soil properties and R^2 and p-value of the simple linear regression for soil properties with r>0.6.

	r	**R^2**	**p-value**
Iron	0.73	0.74	<0.001
Calcium	0.60	0.53	<0.001



(a) Scatter plot of HCPH versus iron



(b) Scatter plot of HCPH versus calcium

Figure 26 – Scatter plots of HCPH vs soil properties ((a) Iron ; (b) Calcium).

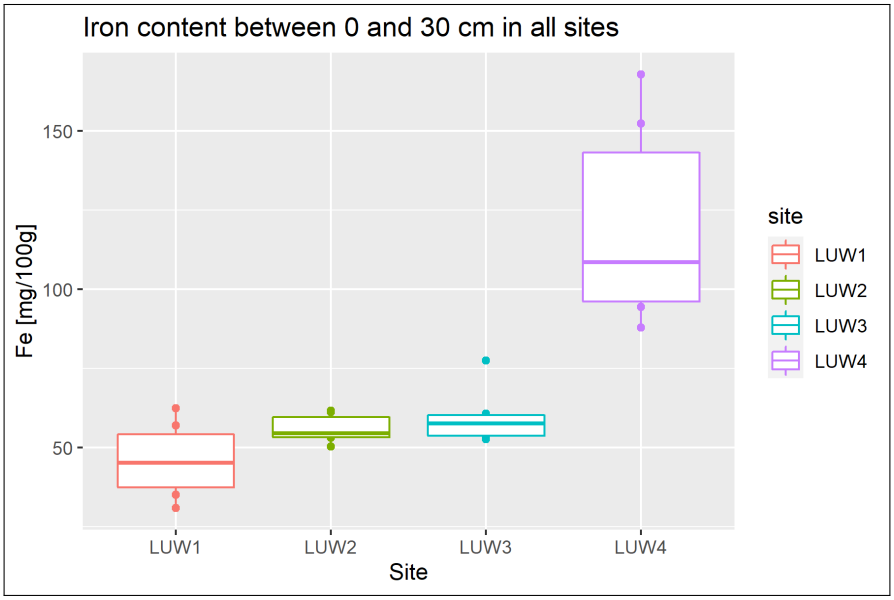
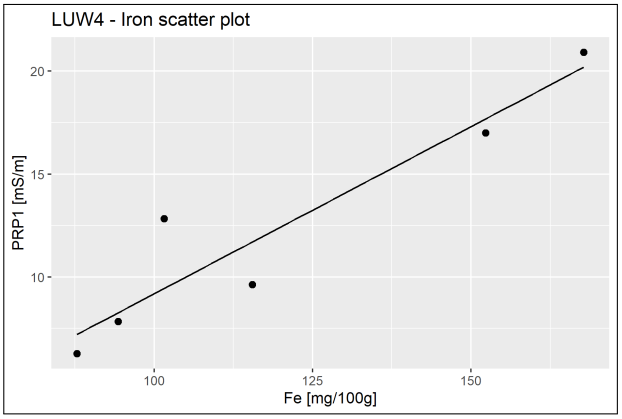
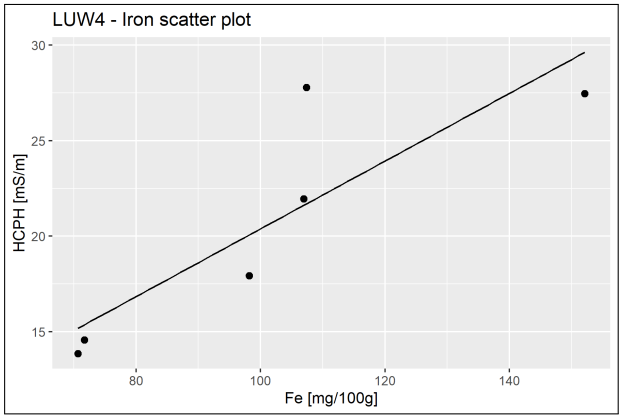


Figure 27 – Boxplot of iron content between 0 and 30 cm for all samples for each site.



(a) LUW4 - Scatter plot of PRP1 versus iron



(b) LUW4 - Scatter plot of HCPH versus iron

Figure 28 – LUW4 - Scatter plots of EMI signal vs iron ((a) PRP1 ; (b) HCPH).

Correlations are clearly higher when LUW4 is not taking into account (tables 7 and 8).

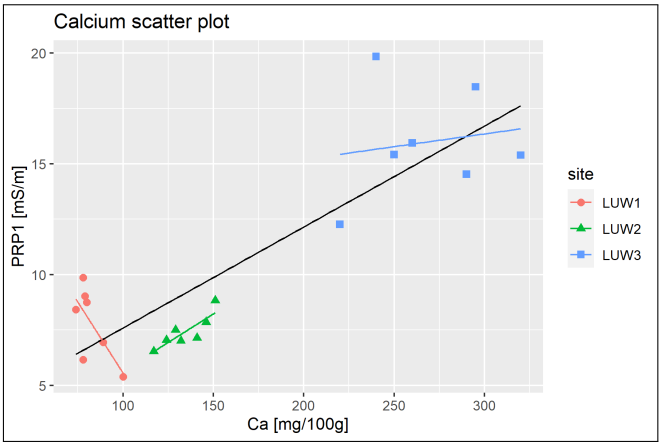
In order to be able to analyse the further management areas for LUW1, LUW2 and LUW3, we will focus on only four properties. Indeed, for most soil properties the link with EC_a isn't strong enough. For instance, for some properties, even if correlations are in the direction of a link between this characteristic and the electrical conductivity and that it is possible to make a significant regression between them, it seems to be only thanks to one of the sites while the two others have no clear tendencies for this properties. Moreover, the soil properties are obviously not independent of each other. It would not be relevant to explain the electrical conductivity by the clay content, the silt content and the sand content at the same time. It was therefore decided to analyse the EMI data and the resulting management areas through the prism of calcium (figure 29), clay (figure 30) and soil water content (figure 31). With the purpose of validating the choice of these three properties, a multiple linear regression was performed for HCPH and PRP1 and these regressions being significant (both p-values < 0.001), these characteristics were kept. In addition to their significance in this data set, clay, soil moisture and calcium have already been linked to the EMI signal in lots of studies (Doolittle and Brevik, 2014).

Table 7 – LUW1, LUW2, LUW3 - Correlations between soil EC_a of PRP1 coil geometry (DOI = 40 cm) and soil properties and R^2 and p-value of the simple linear regression for soil properties with r>0.6.

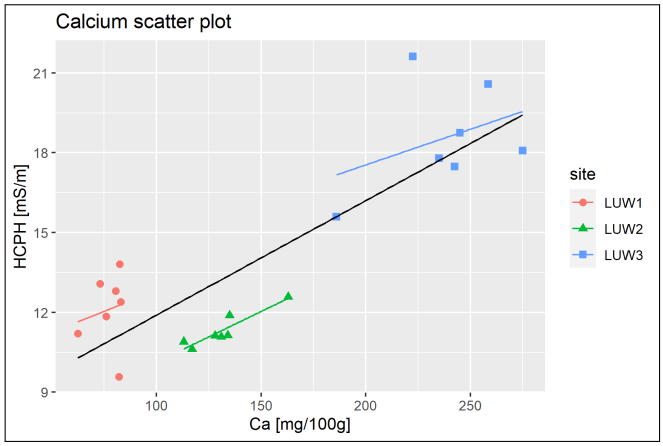
	r	R^2	p-value
Coarse sand	-0.90	0.81	<0.001
Calcium	0.86	0.74	<0.001
Clay	0.84	0.71	<0.001
CEC	0.82	0.68	<0.001
pH H20	0.79	0.55	<0.001
pH KCl	0.77	0.51	<0.001
Moisture content	0.76	0.55	<0.001
Nitrogen	0.64	0.49	<0.001
Fine silt	0.62	0.38	0.003

Table 8 – LUW1, LUW2, LUW3 - Correlations between soil EC_a of HCPH coil geometry (DOI = 65 cm) and soil properties and R^2 and p-value of the simple linear regression for soil properties with r>0.6.

	r	R^2	p-value
Coarse sand	-0.93	0.87	<0.001
Calcium	0.84	0.72	<0.001
Moisture content	0.83	0.69	<0.001
Clay	0.83	0.68	<0.001
CEC	0.80	0.64	<0.001
pH H20	0.77	0.55	<0.001
pH KCl	0.66	0.42	0.002

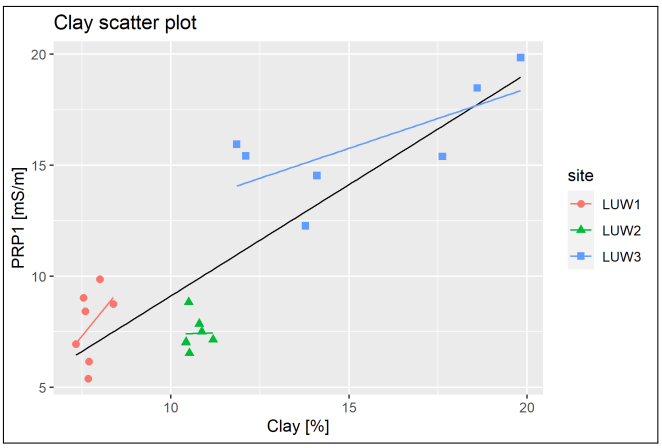


(a) LUW1, LUW2 and LUW3 - Scatter plot of PRP1 vs calcium

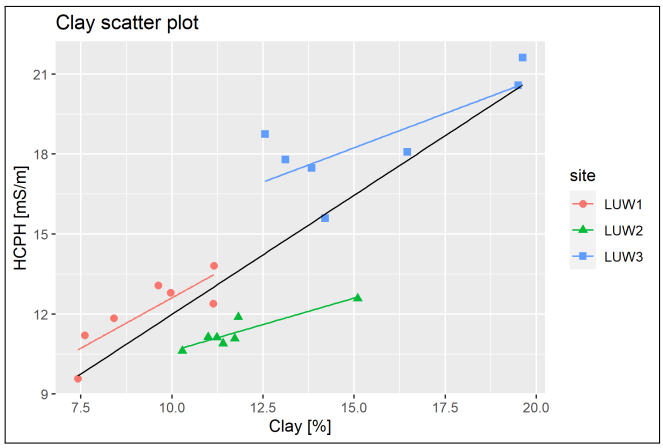


(b) LUW1, LUW2 and LUW3 - Scatter plot of HCPH vs calcium

Figure 29 – LUW1, LUW2 and LUW3 - Scatter plots of EMI signal vs calcium ((a) PRP1 ; (b) HCPH).

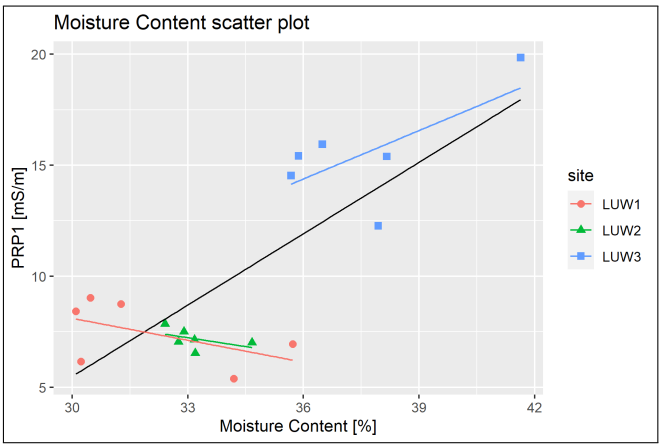


(a) LUW1, LUW2 and LUW3 - Scatter plot of PRP1 vs clay

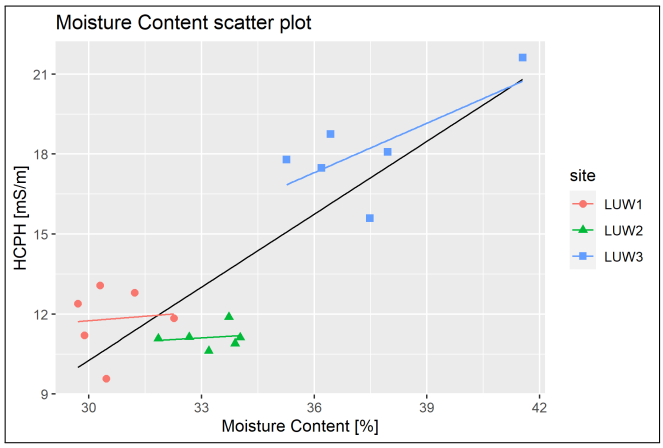


(b) LUW1, LUW2 and LUW3 - Scatter plot of HCPH vs clay

Figure 30 – LUW1, LUW2 and LUW3 - Scatter plots of EMI signal vs clay ((a) PRP1 ; (b) HCPH).



(a) LUW1, LUW2 and LUW3 - Scatter plot of PRP1 vs moisture content



(b) LUW1, LUW2 and LUW3 - Scatter plot of HCPH vs moisture content

Figure 31 – LUW1, LUW2 and LUW3 - Scatter plots of EMI signal vs moisture content ((a) PRP1 ; (b) HCPH).

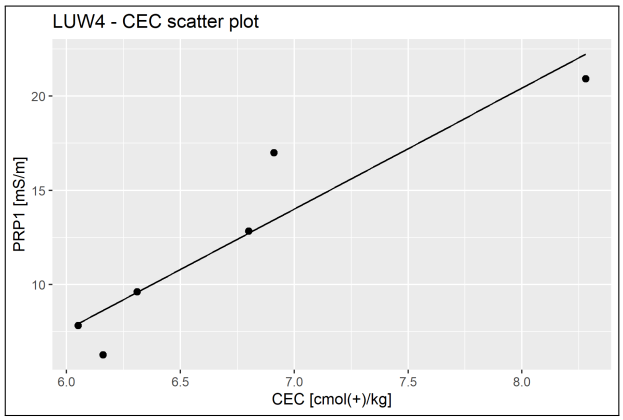
Regarding LUW4, in addition to iron, the choice has been made to focus on iron, CEC, moisture content, SOC and pH for the 30 first cm of the soil (figure 32). Statistics for these soil properties are given in table 9. For the 30-60 cm samples, there are less soil properties which explain significantly the EC_a. However, pH again seems to explain it in addition to iron (figure 33, table 10).

Table 9 – LUW4 - Correlations between soil EC_a of PRP1 coil geometry (DOI = 40 cm) and soil properties and R^2 and p-value of the simple linear regression for selected soil properties.

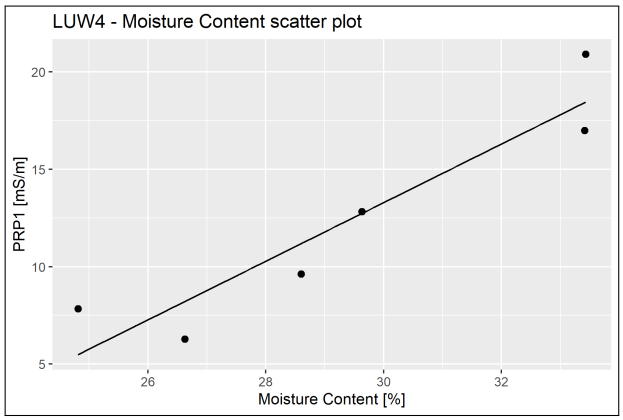
	r	R^2	p-value
Iron	0.94	0.89	0.005
Moisture content	0.94	0.88	0.006
CEC	0.94	0.88	0.006
SOC	0.92	0.85	0.009
pH KCl	0.90	0.81	0.015

Table 10 – LUW4 - Correlations between soil EC_a of HCPH coil geometry (DOI = 65 cm) and soil properties and R^2 and p-value of the simple linear regression for selected soil properties.

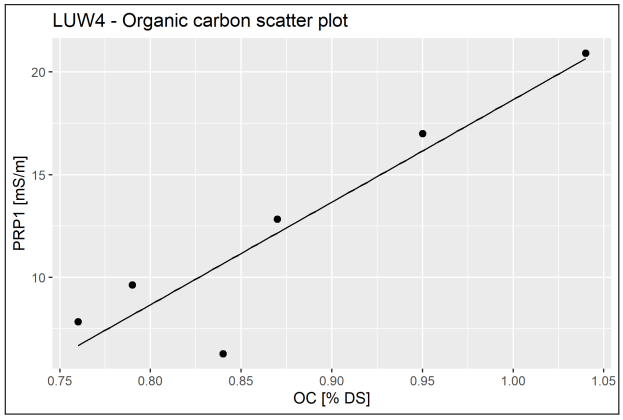
	r	R^2	p-value
pH KCl	0.87	0.76	0.024
Iron	0.86	0.74	0.027



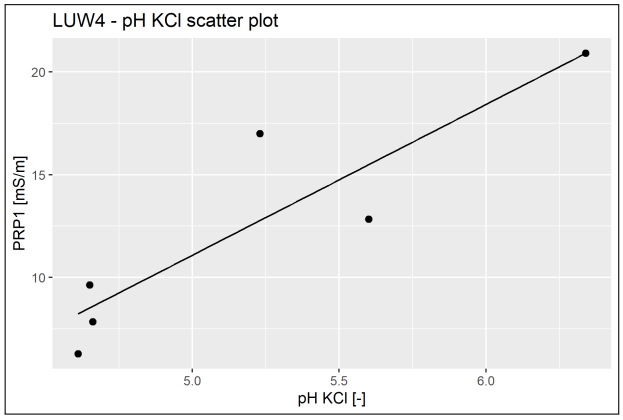
(a) LUW4 - Scatter plot of PRP1 vs CEC



(b) LUW4 - Scatter plot of PRP1 vs moisture content



(c) LUW4 - Scatter plot of PRP1 vs soil organic carbon



(d) LUW4 - Scatter plot of PRP1 vs pH KCl

Figure 32 – LUW4 - Scatter plots of PRP1 vs soil properties ((a) CEC ; (b) Moisture content ; (c) SOC ; (d) pH KCl).

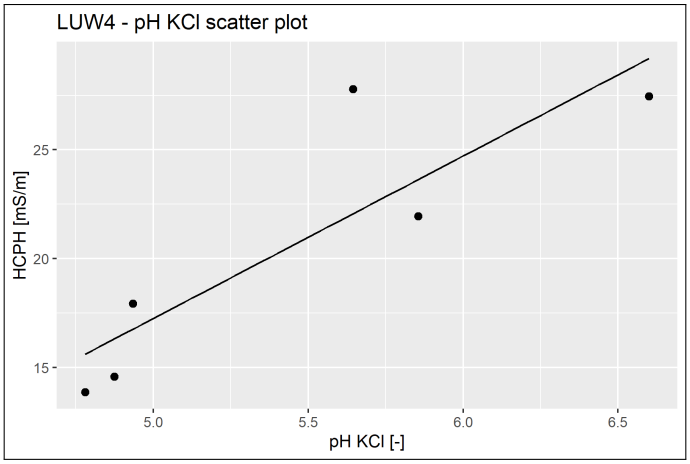


Figure 33 – LUW4 - Scatter plot of HCPH versus pH KCl.

4.3 Clustering analysis

The analysis of the relation between EC_a and soil properties allows for a more precise analysis of the clustering in the different fields. A summary of the different classifications and samples locations is presented in figure 34.

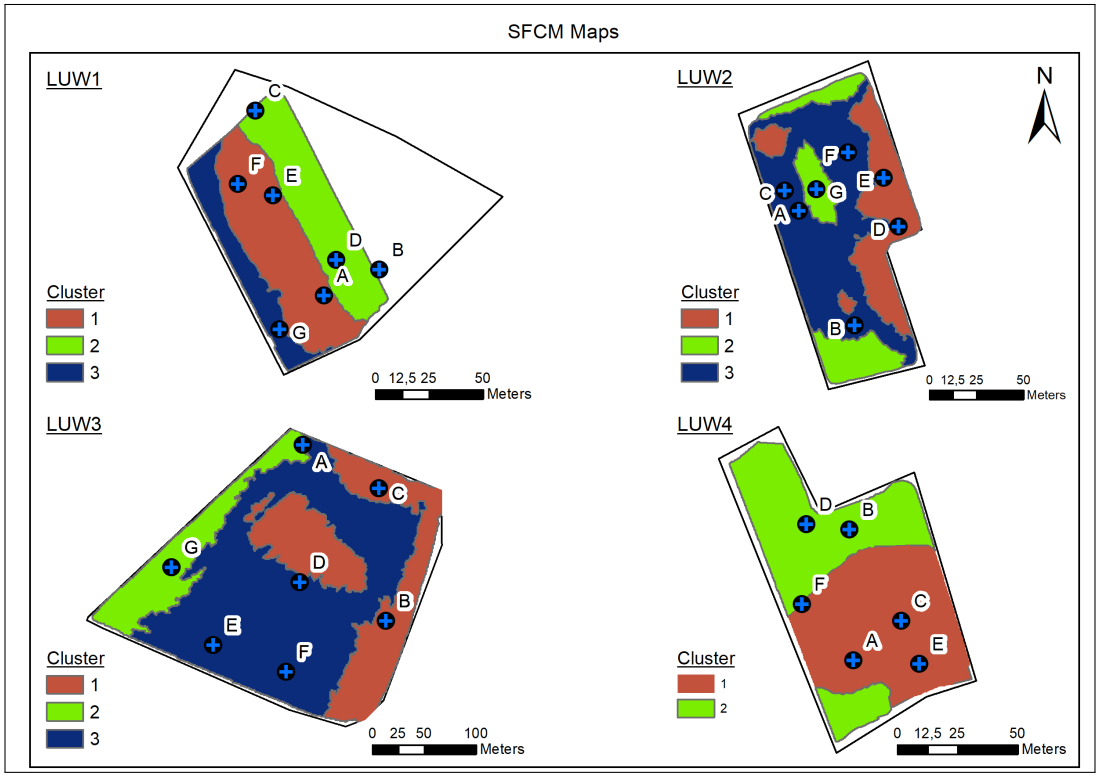


Figure 34 – Clustered EMI maps - summary.

4.3.1 LUW1

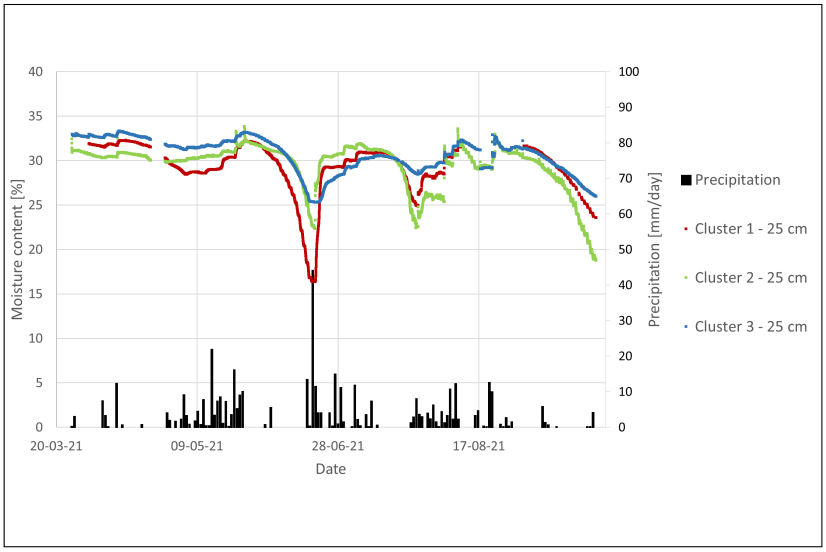
Table 11 presents the mean value and the standard deviation of the EC_a for each cluster and each coils configuration while table 12 shows the results of the soil analysis for the selected soil properties for each samples with the indication of their cluster for LUW1. EC_a is higher, whatever the volume of soil explored, in cluster 3 followed by cluster 1 and finally cluster 2 (table 11). As explained before, it is difficult to explain the differences in cluster 2 from the other two because the topsoil was tilled unlike the rest of the field. However, the samples taken between 30 and 60 cm, and therefore below the ploughed zone, indicate a lower clay content as well as, and this is related, a lower water content than in cluster 1. There is also less calcium. Concerning cluster 3, there is no water content for the single sample which is located in this area. The reason is simple, G sample was collected several months after the EMI scan so the water content couldn't be compared to the EC_a. Indeed, water content is temporary dependant. However, one Sentek probe is located in each of the cluster and shows that the water content is more stable and reacts less quickly to weather variations in the third cluster at 25 cm depth (figure 35). The clay content which is higher than in cluster 2 goes in the same direction. Moreover, the field observation, by installing the Sentek probes, also showed that the area was more humid. Finally, cluster 1 and 3 used to belong to a different field than cluster 2 in the past and that may partly explain the differences between the areas.

Table 11 – LUW1 - EC_a statistics in each cluster.

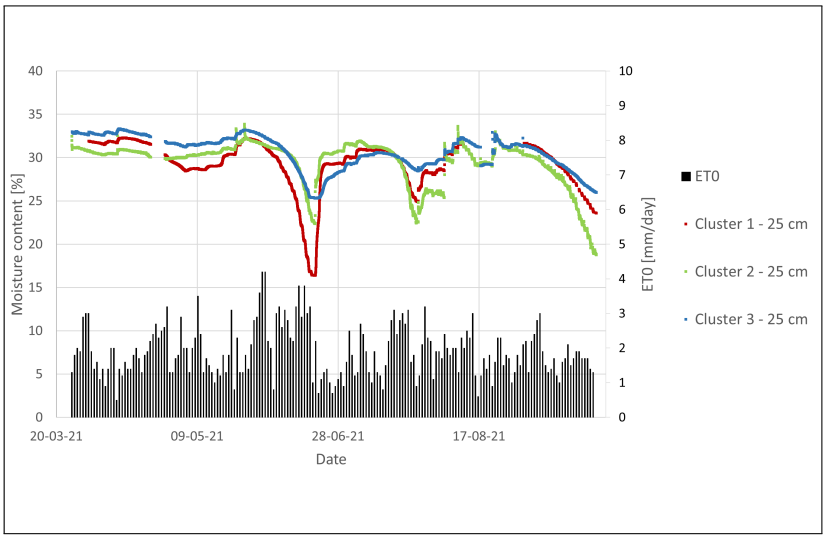
		Cluster		
		1	2	3
PRPH	Mean ECa [mS/m]	5.7	3.8	7.0
	STD [mS/m]	0.5	0.5	1.0
PRP1	Mean ECa [mS/m]	8.9	6.5	10.7
	STD [mS/m]	0.6	0.6	1.3
PRP2	Mean ECa [mS/m]	13.0	10.8	14.8
	STD [mS/m]	0.6	0.9	1.3
HCPH	Mean ECa [mS/m]	12.8	11.3	14.3
	STD [mS/m]	0.6	0.6	1.0
HCP1	Mean ECa [mS/m]	18.0	16.4	19.7
	STD [mS/m]	0.6	0.9	0.9

Table 12 – LUW1 - results of the soil analysis for the selected soil properties.

	Sample	**Cluster**	**Moisture Content [%]**	**Clay [%]**	**Ca [mg/100g]**
0 - 30 cm	A	1	31.27	8.38	80
	B	2	35.72	7.33	89
	C	2	34.19	7.67	100
	D	2	30.23	7.70	78
	E	1	30.10	7.60	74
	F	1	30.48	7.54	79
	G	3	/	8.00	78
30 - 60 cm	A	1	29.33	10.86	66
	B	2	28.81	9.47	63
	C	2	26.74	7.17	64
	D	2	29.53	7.52	47
	E	1	29.33	14.66	92
	F	1	31.96	12.39	82
	G	3	/	14.30	87



(a) Moisture content at 25 cm and daily precipitation



(b) Moisture content at 25 cm and daily ET_0

Figure 35 – LUW1 - Moisture content at 25 cm in each cluster ((a) with daily precipitation ; (b) with daily ET_0).

4.3.2 LUW2

For LUW2, by looking at the mean EC_a of the clusters (table 13) we see that the differences between the clusters are less pronounced than for LUW1. Furthermore, concerning the six first samples, the field could seem very homogeneous. However, sample G, taken after the SFCM was performed, reveals a clear difference in texture which might explain the higher EC_a of the second cluster. This shows the interest of having performed a clustering of the EMI data. If we focus only on the experimental part, near and in which the samples were taken, it is then important to note that there is certainly a zone significantly more clayey in depth corresponding to the cluster 2 and so to the second walnut tree row of the LUW experiment. This could affect the performances of the walnut trees in the future and must therefore be known when analysing further results. The clay layer can lead to poor drainage which is unfavourable to the good development of the walnut tree (Wertheim, 1981; Germain et al., 1999).

Table 13 – LUW2 - EC_a statistics in each cluster.

		Cluster		
		1	2	3
PRPH	Mean EC_a [mS/m]	4.4	5.8	4.8
	STD [mS/m]	0.5	0.8	0.6
PRP1	Mean EC_a [mS/m]	6.5	8.8	7.4
	STD [mS/m]	0.5	0.7	0.4
PRP2	Mean EC_a [mS/m]	9.2	12.0	10.6
	STD [mS/m]	0.7	0.7	0.6
HCPH	Mean EC_a [mS/m]	10.4	12.5	11.3
	STD [mS/m]	0.7	0.7	0.6
HCP1	Mean EC_a [mS/m]	12.9	15.5	14.2
	STD [mS/m]	0.6	0.7	0.6

Table 14 – LUW2 - results of the soil analysis for the selected soil properties.

	Sample	Cluster	Moisture Content [%]	Clay [%]	Ca [mg/100g]
0 - 30 cm	A	3	34.66	10.41	132
	B	3	33.18	11.18	141
	C	3	32.90	10.86	129
	D	1	32.76	10.42	124
	E	1	33.20	10.52	117
	F	3	32.41	10.79	146
	G	2	/	10.50	151
30 - 60 cm	A	3	33.39	12.05	124
	B	3	32.18	10.79	127
	C	3	30.79	12.59	133
	D	1	33.62	10.14	110
	E	1	34.59	12.29	109
	F	3	35.04	12.84	124
	G	2	/	19.70	175

4.3.3 LUW3

In LUW3, the second cluster is the one with the highest EC_a while cluster 1 has the lowest EC_a. Indeed, cluster 2 has more clay. Moisture content is at the highest for A sample which is in the cluster 2 but it is difficult to give an estimation for G sample because on the one hand the clay content is high but on the other hand the elevation on the field is higher than for the other sampling points. It would not be rigorous to draw much other information from the samples because trends are not clear in this case. Nevertheless, as said before, the clusters follow the topography. A silty and organic-rich zone at the bottom of the slope, which corresponds to the cluster 1, can therefore be expected.

Table 15 – LUW3 - EC_a statistics in each cluster.

		Cluster		
		1	2	3
PRPH	Mean ECa [mS/m]	8.3	11.1	9.7
	STD [mS/m]	1.0	1.2	0.9
PRP1	Mean ECa [mS/m]	13.7	18.8	16.2
	STD [mS/m]	1.3	1.7	1.1
PRP2	Mean ECa [mS/m]	19.9	26.5	23.3
	STD [mS/m]	1.6	2.2	1.1
HCPH	Mean ECa [mS/m]	16.3	20.8	18.7
	STD [mS/m]	1.2	1.6	0.9
HCP1	Mean ECa [mS/m]	25.2	31.7	28.6
	STD [mS/m]	1.6	2.4	1.1

Table 16 – LUW3 - results of the soil analysis for the selected soil properties.

	Sample	Cluster	Moisture Content [%]	Clay [%]	Ca [mg/100g]
0 - 30 cm	A	2	41.64	19.82	240
	B	1	35.68	14.10	290
	C	1	37.94	13.78	220
	D	3	38.17	17.62	320
	E	3	36.50	11.85	260
	F	3	35.87	12.10	250
	G	2	/	18.60	295
30 - 60 cm	A	2	41.46	19.43	205
	B	1	36.71	13.57	195
	C	1	37.03	14.63	152
	D	3	37.74	15.29	230
	E	3	36.37	13.26	230
	F	3	34.65	14.13	220
	G	2	/	20.40	222

4.3.4 LUW4

Concerning LUW4, table 17 shows that the cluster 2 have a clearly higher EC_a than the cluster 1. The soil analysis (table 18) confirms the EMI results (figure 23) and the soil map (figure 10). Indeed, the samples taken in the second cluster have a higher water content which is explained by the poor drainage and can explain the EC_a. The cluster 2 is therefore a wetter area, with higher organic carbon content, higher pH and higher iron content. Sample F isn't taken into account here because it is at the frontier between the two zones. It can not be seen with the soil analysis but the soil in the cluster 2 (figures 36 and 37) have a white colour synonymous with gleying while the soil in cluster 1 has colours that go more towards orange (figure 38). The water table level seems therefore to be higher in cluster 2. As far as the expected performances of the trees are concerned, the poor drainage area represented by cluster 2 could significantly affect the trees in the future, excessive humidity and insufficient aeration of the soil being one of the most impacting elements for the walnut tree (Wertheim, 1981; Germain et al., 1999). However, pH is very low in cluster 1 and it could also affect the performance of the trees (Wertheim, 1981; Germain et al., 1999).

Table 17 – LUW4 - EC_a statistics in each cluster.

		Cluster	
		1	2
PRPH	Mean ECa [mS/m]	4.5	9.1
	STD [mS/m]	0.9	2.1
PRP1	Mean ECa [mS/m]	9.1	17.8
	STD [mS/m]	1.8	3.0
PRP2	Mean ECa [mS/m]	19.3	32.8
	STD [mS/m]	3.7	4.1
HCPH	Mean ECa [mS/m]	18.5	25.4
	STD [mS/m]	2.3	2.2
HCP1	Mean ECa [mS/m]	31.1	44.4
	STD [mS/m]	4.3	3.9

Table 18 – LUW4 - results of the soil analysis for the selected soil properties.

	Sample	Cluster	CEC [cmol(+)/kg]	Moisture content [%]	Organic carbon [%]	pH KCl [-]	Iron [mg/100g]
0-30 cm	A	1	6.05	24.81	0.76	4.66	94.3
	B	2	6.91	33.41	0.95	5.23	152.3
	C	1	6.16	26.63	0.84	4.61	87.8
	D	2	8.28	33.43	1.04	6.34	167.8
	E	1	6.31	28.60	0.79	4.65	115.5
30-60 cm	A	1				5.09	49.1
	B	2				6.06	62.5
	C	1				4.95	53.5
	D	2				6.86	136.4
	E	1				5.22	80.9

Figure 36 – Soil profile of sample B in cluster 2.

Figure 37 – Soil profile of sample D in cluster 2.

Figure 38 – Soil profile of sample C in cluster 2.

4.4 Clustering with NDVI data

SFCM algorithm has been computed with PRPH, PRP1 and NDVI data (figure 39) to compare with the results of clustering with all the EMI data but without NDVI (figure 40).

For LUW1, the analysis of the different classification index shows that a fourth group was needed (in yellow). Except for this new group, the clusters have remained almost unchanged.

Concerning LUW2, the important element is the disappearance of the zone following the clayey area. This area didn't therefore impact the vegetation index. Expectations about low walnut tree performance might therefore be discussed if crops are not affected. However, it is still early to predict whether the trees will develop well or not and if differences are observed in the future in this area, this clustering could be used to justify these differences.

No clear differences are observed in LUW3. What could be considered as edges effect in the Western part of the field for the NDVI (figure 39) goes in fact in the same direction as EMI data and the a cluster following this area is observed in both classifications.

Finally, concerning LUW4, lower plants performances is confirmed in the Northern part of the field (figure 39), with poor drainage and high water table level. This confirms that attention must be paid to the health of the trees in this area.

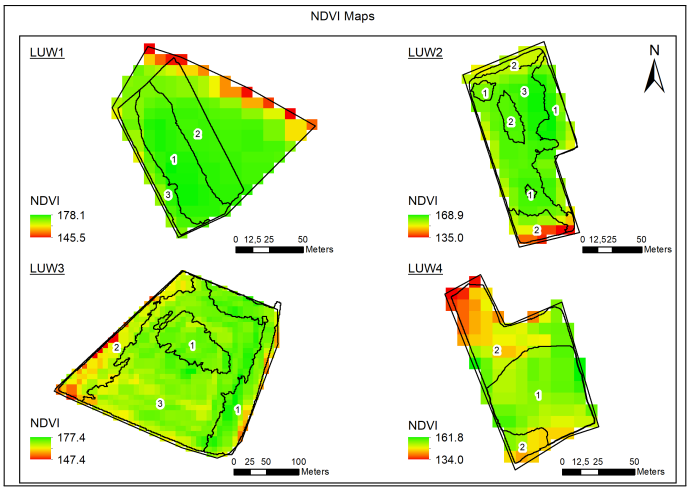
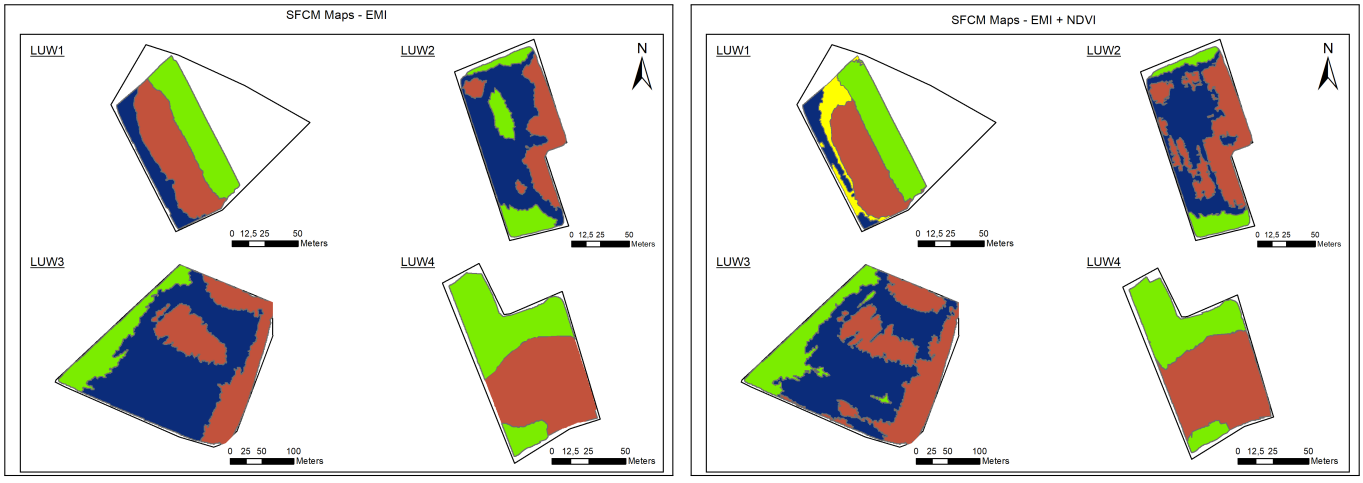


Figure 39 – NDVI maps.



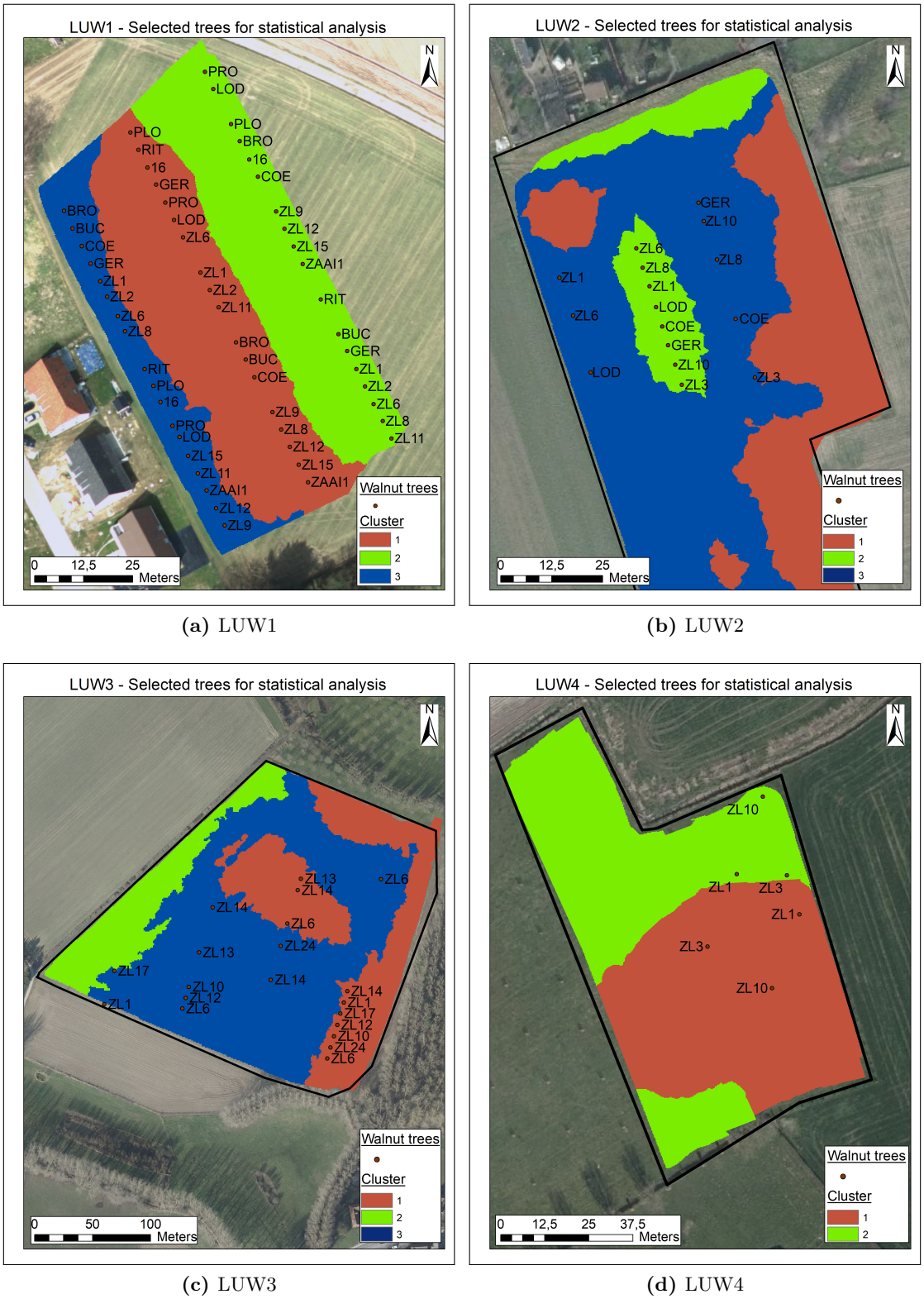
(a) Clustered maps : EMI

(b) Clustered maps : EMI + NDVI

Figure 40 – Clustered maps ((a) EMI ; (b) EMI + NDVI).

4.5 First results of tree characteristics

In order to have comparable groups, the first step was to select trees in each cluster (figure 41). A variety had to be represented an equal number of times in each group for a site. It is important to remember that the clusters determined previously do not take into account the experimental design of the sites and that it is therefore not appropriate to select all the trees of a cluster.



(a) LUW1 **(b)** LUW2

(c) LUW3 **(d)** LUW4

Figure 41 – Selected trees for the statistical analysis ((a) LUW1 ; (b) LUW2 ; (c) LUW3 ; (d) LUW4).

Most groups do not follow a normal distribution (table 19), which is one of the assumption of the ANOVA. The results must therefore be considered with caution. Regarding the differences in performance in terms of budding time and leaf damages, almost no significant effect is observed (table 19) except for the leaf damage measurement for LUW1 (the second measurement) and LUW3 (p-value = 0.000182 and 0.0362 respectively).

The Tukey's test for LUW1 shows that the walnut trees in the second cluster have a significantly higher damage score than those in the first cluster (p-value = 0.00147) and than those in the third cluster (p-value = 0.00043) (table 20). However, the groups do not follow a normal distribution and the results are then less reliable and moreover the clusters follow the tree rows and therefore the higher damage score in the second cluster may be the effect of proximity between trees that increases the chances of disease transmission or simply the position in the field rather than the effect of spatial heterogeneity of the soil. Furthermore, due to differences in terms of soil management of this second cluster, it is the determined homogeneous soil area for which there is the least certainty as to its relevance.

Concerning LUW3, the Tukey's test shows that the damage score is higher in cluster 3 than in cluster 2 (p-value = 0.0362) (table 20). Once again, the groups do not follow a normal distribution and the results are therefore less reliable, especially when taking into account that the p-value is close to not be significant.

Table 19 – Summary of the statistical analyses.

Site	**Variable**	**Shapiro-Wilk test**			**Levene's test**		**One-way ANOVA**	
			p-value	Normal distribution	p-value	Homogeneity of variances	p-value	Significantly different
LUW1	Budding time	Cluster 1	0.000276	NO				
		Cluster 2	0.000720	NO	0.84	YES	0.999	NO
		Cluster 3	0.00341	NO				
	Leaf damages 1	Cluster 1	0.000489	NO				
		Cluster 2	0.000804	NO	0.05984	YES	0.14	NO
		Cluster 3	0.000000321	NO				
	Leaf damages 2	Cluster 1	0.00684	NO				
		Cluster 2	0.0318	NO	0.07001	YES	0.000182	YES
		Cluster 3	0.0000299	NO				
LUW2	Budding time	Cluster 2	0.977	YES	0.8798	YES	0.936	NO
		Cluster 3	0.877	YES				
	Leaf damages 1	Cluster 2	0.000892	NO	0.5057	YES	0.303	NO
		Cluster 3	0.0672	YES				
	Leaf damages 2	Cluster 2	0.000165	NO	0.642	YES	0.705	NO
		Cluster 3	0.000892	NO				
LUW3	Budding time	Cluster 1	0.000486	NO	0.2934	YES	0.686	NO
		Cluster 3	0.637	YES				
	Leaf damages	Cluster 1	0.0391	NO	0.4442	YES	0.0362	YES
		Cluster 3	0.0548	YES				
LUW4	Budding time	Cluster 1	0.637	YES	1	YES	0.252	NO
		Cluster 2	0.637	YES				
	Leaf damages	Cluster 1	0	NO	0.5614	YES	0.768	NO
		Cluster 2	0	NO				

Table 20 – Results of Tukey's test.

Site	Variable	Tukey's test					
		Group		Estimate difference	Lower confidence (95 %)	Upper confidence (95 %)	p-value
LUW1	Leaf damages 2	Cluster 1	Cluster 2	1.06	0.369	1.74	0.00147
		Cluster 1	Cluster 3	-0.111	-0.798	0.576	0.92
		Cluster 2	Cluster 3	-1.17	-1.85	-0.48	0.00043
LUW3	Leaf damages	Cluster 1	Cluster 3	-1	-1.93	-0.0727	0.0362

Although significant differences were observed in two of the sites, the experiment is still too recent and the trees too young to draw any conclusions. However, if these differences continue in the future of the experiment, conclusions could be made.

Finally, this is outside the field of statistics, but it is important to note that for LUW4, out of 4 trees that died during 2021, two are located in cluster 2 (which represents 6 trees out of 22 in total) and another is located almost at the border of this cluster (figure 42). One hypothesis is that since 2021 was a very wet year and cluster 2 of LUW4 was also particularly wet, the trees in this area suffered from the lack of air for roots.

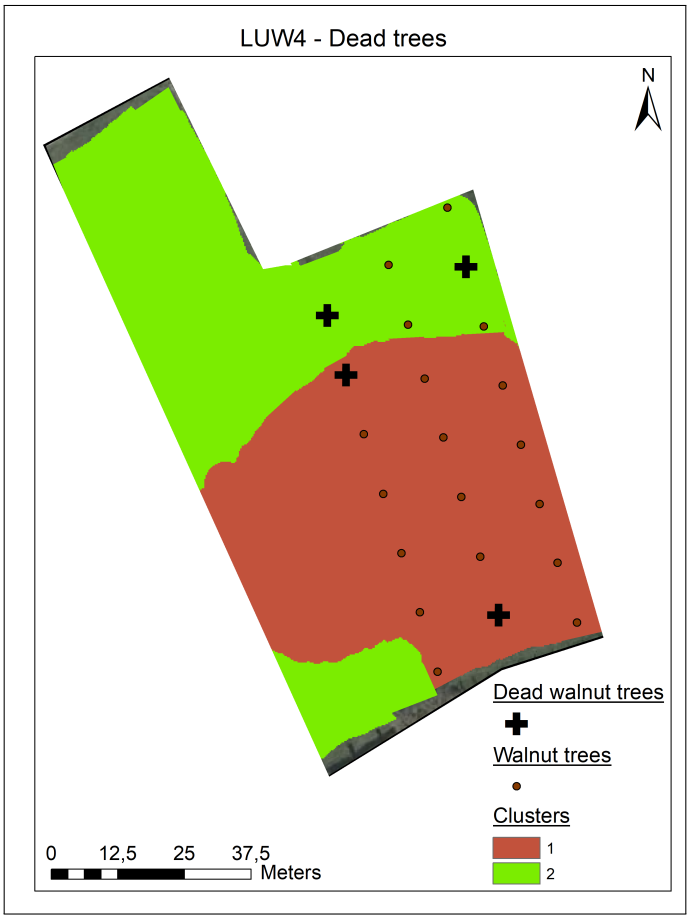


Figure 42 – LUW4 - Dead walnut trees.

5 Discussion

5.1 Soil mapping and determination of homogeneous zones

5.1.1 EMI soil scanning

An EMI soil scanning was performed in March 2021 in the four studied AF sites. This scanning provided six soil EC_a maps per field based on the six different coil orientations : horizontal coplanar dipole mode with a coil spacing of 0.5, 1 and 2 m (HCPH, HCP1 and HCP2) and perpendicular dipole mode with a coil spacing of 0.6, 1.1 and 2.1 m (PRPH, PRP1 and PRP2). For LUW2 and LUW3, the soil scan went well. However, concerning LUW1 and LUW4, the data need to be discussed.

For LUW1, the land cover and soil management were not homogeneous when the EMI scanning was performed. This biased the results in two ways. Firstly, the EMI rasters had to be clipped to eliminate one of the three land cover of the dataset and to only focus on the two within the long-term experimental trial. However, this clipped part nowadays also belongs to the experiment and it would have been interesting to take it into account in the search for soil homogeneous areas. Furthermore, this would have added heterogeneity to the data and may have resulted in areas with larger differences or in more areas. Then the fact that the field had different land covers within the agroforestry system and that the rasters couldn't therefore be clipped except by not considering one of the rows of trees biased the results of the clustering. Indeed, the ploughing of the field had a important weight in the results of the soil scan and one of the three clusters directly follows the ploughed portion of the plot. To focus on the deeper layers in the scan, so as to eliminate heterogeneity caused by recent tillage in the top layer, was not a solution because the impact of the first 30 cm of soil was important for all coil configurations. The soil scanning should have been performed on a homogeneous field especially in this case in which the objective is to search for homogeneous soil areas.

Regarding LUW4, some metal fences were around the trees to protect them and the different rasters were then clipped because the metal distorted the scan results. However, the impact of the metal fences was only felt around the trees and the rest of the field had reliable data.

EMI soil scanning has been done only once on each field. Therefore, it does not take into account a potential temporal heterogeneity of the sites. Indeed, soil moisture has a huge impact on the soil EC_a and soil moisture on a field is time dependent. Additional soil scans could have brought information about the temporal heterogeneity.

5.1.2 Data clustering

Several soil zones considered as homogeneous have been determined by EMI data clustering with a spatial fuzzy c-means algorithm in each sites.

The clustering was performed based on five EMI data rasters and then compared with a clustering including NDVI data. The NDVI datasets used had a resolution of 8 m x 8 m which is low for fields that are sometimes less than 1 ha in surface area.. That is why these datasets were used only in addition to compare with the previously determined areas instead of being included since the beginning. The use of high resolution NDVI, obtained for instance with a drone, could add a reliable information on the plant performances in the field. This would directly echo the reason why clusters are determined : to determine homogeneous areas in order to interpret the performance of trees.

5.2 Relationships between EMI signal and soil properties

In each sites, six soil samples have been taken between 0 and 30 cm and six others between 30 and 60 cm to analyse the soil. The sample locations were determined using a conditioned Latin hypercube algorithm based on the EMI signal. The goal was to optimally sample the variability in the EM data while limiting

the number of samples so as not to increase the cost for soil analyses compared to a simple cross-sampling. This objective is partly achieved. However, this led to two issues.

Firstly, relying on the cLHS method to determine the sampling location and determining the clusters by another algorithm, the SFCM, led to the need to resample in some areas. It was therefore not possible anymore to analyse the soil moisture of these new samples in regard to the EMI data because of the time shift between sampling and scanning. Moreover, the cLHS method showed a failure by not designating a sample in an area that eventually showed great differences in terms of texture in LUW2.

Secondly, and it is the main problem, a set of six samples is too low to determine in a statistically sound way the effect of a soil property on the EC_a. Even when regrouping LUW1, LUW2 and LUW3 datasets, 18 samples is still a quite limited number. Fortunately, for some properties, and consistent with the literature, it was possible to identify a relationship between the EMI signal and the soil analysis results. However, at the field level this was not always the case or worse, the relationship was different. The assumption was made that this was due to the small number of samples, not always representative enough, and that the real relationship is the one determined based the 18 samples. It should be noted, however, that even with more samples collected in each site, it may still have been necessary to combine the data in order to obtain conclusive results. In general, the greater the variability within a measured property, the stronger the correlations with EC_a measurements (Doolittle and Brevik, 2014).

Finally, a lot of properties have been analysed while only a few of them explain the EMI signal. With the benefit of hindsight, it would have been more appropriate to test initially only the properties that are most closely related to soil electrical conductivity in the literature for the locations determined thanks to the cLHS method. Then, once the clusters determined via SFCM, a composite sample could have been taken from each cluster to test the other properties.

5.3 First results of trees performance

A statistical analysis of the budding time and the health of walnut trees in the light of the homogeneous areas was performed. Although significant differences were observed for the health of the leaves, it is more prudent not to draw any conclusions given that the size of the groups tested is limited and that the hypothesis of normality of the populations is not met.

Then, for the other statistical analyses, any significant effects were observed.

However, all these results should be viewed with caution.

Above all, the trees are two years old and were planted in January 2021. They are therefore still very young and are located in the different clusters for less than a year. Results over several years would have allowed clearer conclusions to be drawn.

An other issue in the analysis is that the agroforestry systems have been designed to be analysed in such a way that the walnut varieties are evenly spread over the field and tree rows, so as to be able to consider the trees from the same varieties as independent replicates. In two of the sites (LUW1 and LUW2), the design furthermore took into account the ability to assess two factors : crop performance in the alleys as affected by the phenology of the walnut trees, by planting the commercial varieties and the late budding varieties in separate block. The clusters representing homogeneous soil areas were determined after the plantation and independently of the design. In the case of LUW1, by chance, the clusters follow the tree rows and each walnut variety is represented at least once per row. It is therefore easily feasible to select one representative group of walnut trees per cluster for the statistical analysis. However it is not the case for the other experimental sites in which only two clusters can be analysed. Moreover, the number of trees in some clusters is too small to allow a reliable analysis.

5.4 Soil moisture dynamics

The Sentek probes, used to measure the moisture content in LUW1, were not calibrated. Conclusions could then be made only concerning the relative values. The calibration of the probes could refine the analysis by allowing the use of the absolute values. Moreover, the calibration could make it easier to model the water content and thus be able to estimate the future of the water dynamics in the plot.

5.5 Overview

The purpose of the present study was to characterise the soil conditions and heterogeneity in four new experimental AF sites with late-budding walnut trees in Flanders by developing a methodology for the determination of homogeneous soil zones and to identify the impact of these growth conditions on the tree performance.

For LUW1, the clustering combined with the soil analysis mainly shows an wetter area in the Western part of the field. Clustering in LUW2 revealed an area with a clay layer in depth, which is not present in the rest of the field. In LUW3, the clustering, and therefore the electrical conductivity, follow the topography and the soil types of the field. Finally, LUW4 analysis highlights a very wet area with a water table that rises close to the surface.

A first remark that can be made is that, for LUW1 and LUW2, the clustering method determined three different zones while the global analysis concludes to only one single zone with really different characteristics compared to the two others. This highlights the fact that there is a dichotomy between the high accuracy and resolution of the EMI data and the goal of determining areas of soil considered homogeneous for trees that are in fact quite tolerant to the soil type and in small-sized plots.

However, although some determined soil zones do not actually vary much from others, the methodology used has highlighted other areas that in terms of texture and moisture content may affect the future performance of walnut trees.

In a broader context, there is an interest in replicating this method, while considering the criticisms and adjustments mentioned previously, for other long-term AF experiments. First, it allows for a more detailed analysis of the soil conditions at the beginning of the experiment. Then, it can refine the understanding of the differences in future results of the trees in the experiment. Even if the experimental designs are created to take into account a potential spatial heterogeneity of the soil (randomised blocks, Latin square, etc.), the clusters will refine the understanding of this heterogeneity.

6 Personal contribution of the student

The main goals for the student was to characterise the soil conditions and heterogeneity in four new experimental AF sites by determining soil homogeneous zones. With this purpose, the activities carried out by the student are :

- The collection of the soil samples.

- The soil moisture analysis of the soil samples.

- The installation of the Sentek probes in LUW1.

- The assessment of a leaf damage score in LUW3 and LUW4.

- The choice of an algorithm for the clustering and the realisation of this clustering.

- The development of a innovative method to identify different soil homogeneous area in AF systems.

- All the analyses of the results.

- Interviews of AF key informants which have improved the knowledge of the student about AF.

In each field, at least one soil zone with significant differences has been identified which may help to explain differences in development or performance of walnut trees in the future of this experiment. Moreover, the developed methodology could be applied in other experimental AF sites.

7 Conclusion

In the new AGROFORESTRY 2025 project, seven experimental sites for a long-term trial of agroforestry with late-budding walnut trees have been launched early 2021 in Flanders, of which four were examined in the framework of this master thesis : LUW1, LUW2, LUW3 and LUW4. An electromagnetic induction soil scanning was conducted on each of these sites in March and soil sampling took place. The EMI soil scanning was simultaneously done with six different coil orientations : PRPH, PRP1, PRP2, HCPH, HCP1 and HCP2. The five first orientations, corresponding to a maximum depth of investigation of 1.5 m, have been implemented in a spatial fuzzy c-means algorithm with the purpose of determining homogeneous soil zones within each field.

In each field, at least one zone with significant differences in terms of soil texture, water table level or water content has been identified. Differences in performance could therefore occur between trees in this zone and those in other zones in the future of the experiment. On the other hand, the methodology developed in this study also determined areas between which there is no large difference. Since walnut tree is not very sensitive to small soil differences, these areas seem to be of limited interest.

Nowadays, the analysis of performance of the walnut trees didn't prove with certainty a significant influence of the determined homogeneous soil zones on the trees. However, the trees are still young and were only planted one year ago. Since the determination of zones is intended to explain differences in tree performance over the long term, more conclusive results are expected in the future.

Prospects in order to improve this methodology of determination and description of homogeneous soil zones are possible. Firstly, now that several soil properties have been shown to be related to soil electrical conductivity in the different experimental sites, more soil samples could be collected as a result of a future scan to analyse only those specific properties. Moreover, samples could be taken according to the exact depth of investigation of the different coils geometry. In this way, robust models could possibly be built to define for each site, or for all the sites in once, the relationship between these properties and EC_a. Secondly, the relevance of the determined soil zones will have to be evaluated during the future of the experiment by analysing the results obtained in the light of the different clusters. Finally, as for a broader view of this methodology, it could be applied in other long-term AF experiments in order to provide a better understanding of the initial soil conditions and especially of the spatial heterogeneity of the experimental sites.

Bibliography

1. Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). FAO Irrigation and Drainage Paper No. 56 - Crop Evapotranspiration. (56).

2. Allred, B. J., Ehsani, M. R., and Daniels, J. J. (2008). General considerations for geophysical methods applied to agriculture. *Handbook of Agricultural Geophysics*, pages 3–16.

3. Banel, D. K. and Hu, F. B. (2009). Effects of walnut consumption on blood lipids and other cardiovascular risk factors : a meta-analysis and systematic review. *The American Journal of Clinical Nutrition*, 90(1) :56–63.

4. Bemah, P., Böhm, R., Brünnich, J., Chan, W.-T., Filibeck, J., Hildreth, A., Krebs, M., McCormick, A., Myers, B. E., Schmitz, S., and Schwarzenberg, S. (2012). Export Opportunity Surveys - The Market for Walnuts in Germany and the United States. (August) :39.

5. Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media.

6. Bezdek, J. C., Ehrlich, R., and Full, W. (1984). Fcm : The fuzzy c-means clustering algorithm. *Computers & geosciences*, 10(2-3) :191–203.

7. Boaga, J. (2017). The use of fdem in hydrogeophysics : A review. *Journal of Applied Geophysics*, 139 :36–46.

8. Borremans, L., Reubens, B., Van Gils, B., Baeyens, D., Vandevelde, C., and Wauters, E. (2016). A sociopsychological analysis of agroforestry adoption in Flanders : understanding the discrepancy between conceptual opportunities and actual implementation. *Agroecology and Sustainable Food Systems*, 40(9) :1008–1036.

9. Borremans, L., Reubens, B., and Wauters, E. (2018). How to make agroforestry pay off ? using its values to create economic incentive pathways.

10. Cai, W., Chen, S., and Zhang, D. (2007). Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern Recognition*, 40(3) :825–838.

11. Calamita, G., Perrone, A., Brocca, L., Onorati, B., and Manfreda, S. (2015). Field test of a multi-frequency electromagnetic induction sensor for soil moisture monitoring in southern italy test sites. *Journal of Hydrology*, 529 :316–329.

12. Chuang, K. S., Tzeng, H. L., Chen, S., Wu, J., and Chen, T. J. (2006). Fuzzy c-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics*, 30(1) :9–15.

13. Consortium Agroforestry Vlaanderen (2021). Walnoot (juglans regia en juglans nigra). `https://www.agroforestryvlaanderen.be/nl/nieuws/walnoot`. Accessed on 2021-12-15.

14. Corwin, D. (2008). Past, present, and future trends of soil electrical conductivity measurement using geophysical methods. *Handbook of Agricultural Geophysics*.

15. Corwin, D. and Scudiero, E. (2019). Chapter one - review of soil salinity assessment for agriculture across multiple scales using proximal and/or remote sensors. volume 158 of *Advances in Agronomy*, pages 1–130. Academic Press.

16. Delefortrie, S., De Smedt, P., Saey, T., Van De Vijver, E., and Van Meirvenne, M. (2014). An efficient calibration procedure for correction of drift in EMI survey data. *Journal of Applied Geophysics*, 110 :115–125.

17. Doolittle, J. A. and Brevik, E. C. (2014). The use of electromagnetic induction techniques in soils studies. *Geoderma*, 223-225(1) :33–45.

18. Egnér, H., Riehm, H., and Domingo, W. (1960). Untersuchungen über die chemische bodenanalyse als grundlage für die beurteilung des nährstoffzustandes der böden. ii chemische extraktionsmethoden zur phosphor-und kaliumbestimmung. *Kungliga Lantbrukshögskolans Annaler*, 26 :199–215.

19. Eichhorn, M. P., Paris, P., Herzog, F., Incoll, L. D., Liagre, F., Mantzanas, K., Mayus, M., Moreno, G., Papanastasis, V. P., Pilbeam, D. J., Pisanelli, A., and Dupraz, C. (2006). Silvoarable systems in Europe - Past, present and future prospects. *Agroforestry Systems*, 67(1) :29–50.

20. European Commission (2019). Tackling climate change. `https://ec.europa.eu/info/food-farming-fisheries/sustainability/environmental-sustainability/climate-change_en`. Accessed on 2021-11-27.

21. FAO (2015). Agroforestry : Definition. `http://www.fao.org/forestry/agroforestry/80338/en/`. Accessed on 2021-03-01.

22. FAO (2021). *The State of Food and Agriculture 2021. Making agrifood systems more resilient to shocks and stresses.* Rome, FAO.

23. Gelb, J. and Apparicio, P. (2021). Apport de la classification floue c-means spatiale en géographie : essai de taxinomie socio-résidentielle et environnementale à Lyon. *Cybergeo : European Journal of Geography*.

24. Genot, V., Colinet, G., and Bock, L. (2007). La fertilité des sols agricoles et forestiers en Région wallonne. Technical report.

25. Germain, E., Prunet, J.-P., Garcin, A., et al. (1999). *Le noyer*. Centre Technique Interprofessionnel des Fruits et Légumes (CTIFL).

26. Goktepe, A. B., Altun, S., and Sezer, A. (2005). Soil clustering by fuzzy c-means algorithm. *Advances in Engineering Software*, 36(10) :691–698.

27. Gomiero, T., Pimentel, D., and Paoletti, M. G. (2011). Environmental impact of different agricultural management practices : Conventional vs. organic agriculture. *Critical Reviews in Plant Sciences*, 30(1-2) :95–124.

28. Graves, A. R., Burgess, P. J., Liagre, F., Pisanelli, A., Paris, P., Moreno, G., Bellido, M., Mayus, M., Postma, M., Schindler, B., Mantzanas, K., Papanastasis, V. P., and Dupraz, C. (2009). *Farmer Perceptions of Silvoarable Systems in Seven European Countries*, pages 67–86. Springer Netherlands, Dordrecht.

29. Hanesch, M., Scholger, R., and Dekkers, M. J. (2001). The application of fuzzy C-means cluster analysis and non-linear mapping to a soil data set for the detection of polluted sites. *Physics and Chemistry of the Earth, Part A : Solid Earth and Geodesy*, 26(11-12) :885–891.

30. Hanssens, D., Waegeman, W., Declercq, Y., Dierckx, H., Verschelde, H., and De Smedt, P. (2020). High-resolution surveying with small-loop frequency domain electromagnetic systems : Efficient survey design and adaptive processing. *IEEE Geoscience and Remote Sensing Magazine*, PP.

31. Hendrickx, J. and Kachanoski, R. (2002). *Nonintrusive electromagnetic induction*, pages 1301–1310.

32. Herder, M. D., Burgess, P., Mosquera-losada, M. R., Hartel, T., Upson, M., Viholainen, I., Rosati, A., Corroyer, N., Hermansen, J. E., Mirck, J., Palma, J., Pantera, A., Papanastasis, V., Plieninger, T., and Vityi, A. (2015). Preliminary stratification and quantification of agroforestry in Europe. (April) :55.

33. ILVO, BDB, INAGRO, ABC Eco², and NPW (2019). Agroforestry 2025 : veranderingstrajecten gericht op systeemoptimalisatie en haalbare verdienmodellen.

34. INC - International Nut and Dried Fruit Council (2021). *Nuts Dried Fruits Statistical Yearbook*.

35. Institut de Mathématiques de Bordeaux (2015). pages 1–14.

36. IPCC (2019). *Climate Change and Land : an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*.

37. Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering : A review. *ACM Comput. Surv.*, 31(3) :264–323.

38. Jose, S., Gillespie, A. R., and Pallardy, S. G. (2004). Interspecific interactions in temperate agroforestry. *Agroforestry Systems*, 61 :237–255.

39. Lakanen, E. and Erviö, R. (1971). A comparison of eight extractants for the determination of plant available micronutrients in soils. *Acta Agral. Fenn.*, 123 :223–232.

40. Lin, C., McGraw, R. L., George, M. F., and Garrett, H. E. G. (1999). Shade effects on forage crops with potential in temperate agroforestry practices. *Agroforestry Systems*, 44 :109–119.

41. Martinez, G., Vanderlinden, K., Ordóñez, R., and Muriel, J. L. (2009). Can Apparent Electrical Conductivity Improve the Spatial Characterization of Soil Organic Carbon ? *Vadose Zone Journal*, 8(3) :586–593.

42. Mary, F., Dupraz, C., Delannoy, E., and Liagre, F. (1998). Incorporating agroforestry practices in the management of walnut plantations in dauphiné, france : an analysis of farmers 'motivations. *Agroforestry systems*, 43(1) :243–256.

43. McAdam, J. H., Burgess, P. J., Graves, A. R., Rigueiro-Rodríguez, A., and Mosquera-Losada, M. R. (2009). *Classifications and Functions of Agroforestry Systems in Europe*, pages 21–41. Springer Netherlands, Dordrecht.

44. McBratney, A. and Odeh, I. O. (1997). Application of fuzzy sets in soil science : fuzzy logic, fuzzy measurements and fuzzy decisions. *Geoderma*, 77(2) :85–113. Fuzzy Sets in Soil Science.

45. McClean, S. I. (2003). Data mining and knowledge discovery. In Meyers, R. A., editor, *Encyclopedia of Physical Science and Technology (Third Edition)*, pages 229–246. Academic Press, New York, third edition edition.

46. McNeill, J. D. (1980). Electromagnetic terrain conductivity measurement at low induction numbers.

47. Merelle, F. (1998). *L'analyse de terre aujourd'hui*. GEMAS (Groupement d'Etudes Methodologiques pour l'Analyse des Sols). Nantes.

48. Minasny, B. and McBratney, A. B. (2006). A conditioned latin hypercube method for sampling in the presence of ancillary information. *Computers Geosciences*, 32(9) :1378–1388.

49. Minsley, B. J., Smith, B. D., Hammack, R., Sams, J. I., and Veloski, G. (2012). Calibration and filtering strategies for frequency domain electromagnetic data. *Journal of Applied Geophysics*, 80 :56–66.

50. Monteith, J. L., Moss, C. J., Cooke, G. W., Pirie, N. W., and Bell, G. D. H. (1977). Climate and the efficiency of crop production in britain. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 281(980) :277–294.

51. Nagpal, A., Jatain, A., and Gaur, D. (2013). Review based on data clustering algorithms. In *2013 IEEE Conference on Information Communication Technologies*, pages 298–303.

52. NAIR, P. (1985). Classification of agroforestry systems. *Agroforestry Systems*, 3 :97–128.

53. Oldenburg, D. W. and Li, Y. (1999). Estimating depth of investigation in dc resistivity and ip surveys. *Geophysics*, 64(2) :403–416.

54. Ong, C., Corlett, J., Singh, R., and Black, C. (1991). Above and below ground interactions in agroforestry systems. *Forest Ecology and Management*, 45(1) :45–57. Agroforestry : Principles and Practice.

55. Pal, N. and Bezdek, J. (1995). On cluster validity for the fuzzy c-means model. *IEEE Transactions on Fuzzy Systems*, 3(3) :370–379.

56. Pantera, A., Burgess, et al. (2018). Agroforestry for high value tree systems in Europe. *Agroforestry Systems*, 92 :945–959.

57. Raes, D. and Munoz, G. (2009). The ETo Calculator. *Reference Manual Version*, pages 1–3.

58. Reubens, B., Wauters, E., Coussement, T., Van Daele, S., Van Nieuwenhove, T., Balis, J.-P., Pardon, P., Borremans, L., Nelissen, V., Raman, M., Elsen, A., Mertens, J., Reheul, D., and Verheyen, K. (2019). *Agroforestry in Vlaanderen 2014-2019 : Handvatten na 5 jaar onderzoek en praktijkwerking.*

59. Rigueiro-Rodríguez, A., Fernández-Núñez, E., González-Hernández, P., McAdam, J. H., and Mosquera-Losada, M. R. (2008). Agroforestry Systems in Europe : Productive, Ecological and Social Perspectives. *Agroforestry in Europe*, pages 43–65.

60. Robinson, D. A., Abdu, H., Jones, S. B., Seyfried, M., Lebron, I., and Knight, R. (2008a). Eco-Geophysical Imaging of Watershed-Scale Soil Patterns Links with Plant Community Spatial Patterns. *Vadose Zone Journal*, 7(4) :1132–1138.

61. Robinson, D. A., Binley, A., Crook, N., Day-Lewis, F. D., Ferré, T. P. A., Grauch, V. J. S., Knight, R., Knoll, M. D., Lakshmi, V., Miller, R. D., Nyquist, J. E., Pellerin, L., Singha, K., and Slater, L. D. (2008b). Advancing process-based watershed hydrological research using near-surface geophysics : a vision for, and review of, electrical and magnetic geophysical methods. *Hydrological Processes*, 22 :3604–3635.

62. Rousseeuw, P. J. (1987). Silhouettes : A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20(C) :53–65.

63. Schröter, I., Paasche, H., Doktor, D., Xu, X., Dietrich, P., and Wollschläger, U. (2017). Estimating Soil Moisture Patterns with Remote Sensing and Terrain Data at the Small Catchment Scale. *Vadose Zone Journal*, 16(10).

64. Sentek (2011). Calibration Manual : For Sentek Soil Moisture.

65. Sentek (2021). Enviroscan. `https://sentektechnologies.com/product-range/soil-data-probes/enviroscan/`. Accessed : 2021-11-03.

66. Singh, M., Bhattacharjee, R., Sharma, N., and Verma, A. (2017). An improved xie-beni index for cluster validity measure. In *2017 Fourth International Conference on Image Information Processing (ICIIP)*, pages 1–5.

67. Stetco, A., Zeng, X. J., and Keane, J. (2015). Fuzzy C-means++ : Fuzzy C-means with effective seeding initialization. *Expert Systems with Applications*, 42(21) :7541–7548.

68. Stone, E. and Kalisz, P. (1991). On the maximum extent of tree roots. *Forest Ecology and Management*, 46(1) :59–102.

69. Taylor, R. (2005). Apparent conductivity as an indicator of thickness. `https://dualem.com/acit.htm`. Accessed on 2021-12-10.

70. Torralba, M., Fagerholm, N., Burgess, P. J., Moreno, G., and Plieninger, T. (2016). Do European agroforestry systems enhance biodiversity and ecosystem services ? A meta-analysis. *Agriculture, Ecosystems and Environment*, 230(August) :150–161.

71. Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2) :127–150.

72. Van Noordwijk, M., Duguma, L., Dewi, S., Leimona, B., Catacutan, D., Lusiana, B., Oborn, I., Hairiah, K., Minang, P., Ekadinata, A., Martini, E., Degrande, A., and Prabhu, R. (2019). Agroforestry paradigms. In *Sustainable development through trees on farms : agroforestry in its fifth decade*, pages 1–14. World Agroforestry (ICRAF) Southeast Asia Regional Program., Bogor, Indonesia.

73. Vereecken, H., Huisman, J., Pachepsky, Y., Montzka, C., van der Kruk, J., Bogena, H., Weihermüller, L., Herbst, M., Martinez, G., and Vanderborght, J. (2014). On the spatio-temporal dynamics of soil moisture at the field scale. *Journal of Hydrology*, 516 :76–96.

74. von Hebel, C., Reynaert, S., Pauly, K., Janssens, P., Piccard, I., Vanderborght, J., van der Kruk, J., Vereecken, H., and Garré, S. (2021). Toward high-resolution agronomic soil information and management zones delineated by ground-based electromagnetic induction and aerial drone data. *Vadose Zone Journal*, 20(4) :1–18.

75. Wait, J. R. (1982). *Geo-Electromagnetism*. Academic Press.

76. Wertheim, S. (1981). *De teelt van walnoten*.

77. Wezel, A., Casagrande, M., Celette, F., Vian, J. F., Ferrer, A., and Peigné, J. (2014). Agroecological practices for sustainable agriculture. A review. *Agronomy for Sustainable Development*, 34(1) :1–20.

78. Wu, Q. (2018). 2.07 - gis and remote sensing applications in wetland mapping and monitoring. In Huang, B., editor, *Comprehensive Geographic Information Systems*, pages 140–157. Elsevier, Oxford.

79. Xie, X. L. and Beni, G. (1991). A Validity Measure for Fuzzy Clustering.

80. Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3) :645–678.

81. Zass, R. and Shashua, A. (2005). A unifying approach to hard and probabilistic clustering. In *Tenth IEEE International Conference on Computer Vision*, volume 1, pages 294–301.

82. Șvarț, V. (2018). *Inclusive Walnut Value Chain Development in the Republic of Moldova*. PhD thesis.

Appendices

Table 21 – LUW1 - tree performance.

Cluster	Variety	Budding week	Damage leaf 1	Damage leaf 2
1	PLO	0	1	2
1	RIT	0	2	1
1	16	0	1	1
1	GER	0	1	2
1	PRO	0	1	2
1	LOD	0	1	2
1	ZL6	8	0	3
1	ZL1	4	1	3
1	ZL2	8	0	0
1	ZL11	7	0	1
1	BRO	0	1	1
1	BUC	0	0	3
1	COE	0	1	3
1	ZL9	7	0	0
1	ZL8	9	0	1
1	ZL12	4	0	0
1	ZL15	4	2	1
1	ZAAI1	9	0	2
2	PRO	0	1	2
2	LOD	0	1	3
2	PLO	0	0	1
2	BRO	0	1	3
2	16	0	0	3
2	COE	0	1	3
2	ZL9	7	0	3
2	ZL12	4	0	3
2	ZL15	4	0	2
2	ZAAI1	9	0	2
2	RIT	0	0	3
2	BUC	3	0	2
2	GER	0	0	3
2	ZL1	4	0	3
2	ZL2	8	0	3
2	ZL6	8	0	3
2	ZL8	6	1	3
2	ZL11	7	0	2
3	BRO	0	1	1
3	BUC	1	0	2
3	COE	0	1	2
3	GER	0	1	2
3	ZL1	6	0	0
3	ZL2	9	0	0
3	ZL6	8	0	2
3	ZL8	7	0	1
3	RIT	0	1	3
3	PLO	0	2	2
3	16	0	1	1
3	PRO	0	2	2
3	LOD	1	1	2
3	ZL15	4	1	2
3	ZL11	7	0	1
3	ZAAI1	8	0	1
3	ZL12	4	0	2
3	ZL9	6	0	0

Table 22 – LUW2 - tree performance.

Cluster	Variety	Budding week	Damage leaf 1	Damage leaf 2
2	ZL6	10	0	1
2	ZL8	8	0	3
2	ZL1	7	0	1
2	LOD	1	1	1
2	COE	2	1	2
2	GER	4	1	1
2	ZL10	6	1	1
2	ZL3	5	0	1
3	COE	1	1	2
3	ZL1	8	0	1
3	ZL6	10	0	1
3	LOD	2	2	2
3	GER	4	1	1
3	ZL10	6	1	2
3	ZL8	8	0	1
3	ZL3	5	2	2

Table 23 – LUW3 - tree performance.

Cluster	Variety	Budding week	Damage leaf
1	ZL6	8	3
1	ZL24	6	2
1	ZL10	6	1
1	ZL12	6	1
1	ZL17	6	2
1	ZL1	6	2
1	ZL14	9	1
1	ZL14	9	1
1	ZL13	6	4
3	ZL14	9	1
3	ZL14	8	0
3	ZL6	9	2
3	ZL12	5	2
3	ZL1	5	1
3	ZL10	3	0
3	ZL17	6	1
3	ZL24	7	1
3	ZL13	7	0

Table 24 – LUW4 - tree performance.

Cluster	Variety	Budding week	Damage leaf
1	ZL1	8	2
1	ZL10	7	1
1	ZL3	10	1
2	ZL3	5	3
2	ZL10	7	0
2	ZL1	8	0

Table 25 – LUW1 - total soil analysis.

	Sample	Moisture Content (%)	Bulk density (g/cm³)	pH H2O	Clay (%)	Fine silt (%)	Coarse silt (%)	Total silt (%)	Fine sand (%)	Coarse sand (%)	Total sand (%)	Crusting index	CEC (cmol (+) /kg)
	LUW1 A	31.27	1.64	6.61	8.38	9.91	34.67	44.58	42.48	4.56	47.04	1.47	4.29
	LUW1 B	35.72	1.45	6.83	7.33	9.26	31.65	40.91	47.45	4.31	51.76	1.18	5.58
0 - 30 cm	LUW1 C	34.19	1.44	6.99	7.67	7.29	29.54	36.83	50.54	4.96	55.50	1.13	4.84
	LUW1 D	30.23	1.53	6.81	7.70	8.47	34.25	42.71	45.47	4.12	49.59	1.19	6.59
	LUW1 E	30.10	1.66	6.91	7.60	7.98	29.26	37.23	50.43	4.74	55.17	1.35	6.20
	LUW1 F	30.48	1.54	7.06	7.54	7.54	31.30	38.84	48.78	4.84	53.62	1.32	5.76
	LUW1 G			7.00	8.00	10.70	33.50	44.20	43.20	4.70	47.80	1.47	
	LUW1 A	29.33	1.70	6.96	10.86	11.61	32.58	44.19	41.10	3.85	44.95	2.04	4.79
	LUW1 B	28.81	1.73	6.96	9.47	11.37	37.13	48.49	37.66	4.38	42.04	2.50	4.70
30 - 60 cm	LUW1 C	26.74	1.60	7.14	7.17	7.92	30.56	38.48	49.72	4.63	54.35	2.01	6.81
	LUW1 D	29.53	1.65	6.84	7.52	10.52	36.08	46.60	42.06	3.82	45.88	2.71	5.77
	LUW1 E	29.33	1.64	7.16	14.66	12.03	31.57	43.60	38.63	3.11	41.75	2.11	9.03
	LUW1 F	31.96	1.66	7.04	12.39	12.01	31.90	43.91	40.52	3.18	43.70	1.92	7.67
	LUW1 G			7.40	14.30	12.00	35.30	47.30	35.00	3.40	38.40	2.17	

	Sample	SOC (%)	pH KCl	Total N (%)	Fe (mg/100g)	K (mg/100g)	Mg (mg/100g)	Ca (mg/100g)	Mn (mg/100g)	Na (mg/100g)	P (mg/100g)	Al (mg/100g)
	LUW1 A	1.07	5.84	0.10	57.00	31.00	17.70	80.00	17.74	< 1.92	21.00	25.41
	LUW1 B	1.29	6.06	0.12	30.90	41.00	17.20	89.00	13.87	< 1.92	19.20	21.48
0 - 30 cm	LUW1 C	1.37	6.10	0.11	62.40	32.00	18.00	100.00	19.78	< 1.92	26.00	25.31
	LUW1 D	1.16	6.00	0.11	35.00	35.00	17.60	78.00	13.47	< 1.92	17.40	21.89
	LUW1 E	1.00	5.72	0.09	45.80	27.00	16.00	74.00	11.26	< 1.92	18.80	23.74
	LUW1 F	1.20	5.90	0.11	44.60	35.00	16.00	79.00	12.29	< 1.92	22.00	24.02
	LUW1 G		5.90	0.09		33.00	13.00	78.00			9.00	
	LUW1 A	0.44	5.57	0.05	31.40	25.00	22.00	66.00	5.82	< 1.92	7.70	23.58
	LUW1 B	0.36	5.65	0.04	23.30	22.00	16.90	63.00	4.15	< 1.92	5.60	22.19
30 - 60 cm	LUW1 C	0.63	6.05	0.06	42.90	25.00	19.20	64.00	11.87	< 1.92	11.10	20.27
	LUW1 D	0.33	5.61	0.04	19.30	19.00	15.80	47.00	4.35	< 1.92	4.80	17.12
	LUW1 E	0.22	5.39	0.03	19.80	27.00	28.00	92.00	2.81	2.40	3.80	26.24
	LUW1 F	0.27	5.47	0.04	32.00	29.00	30.00	82.00	3.10	< 1.92	3.10	32.71
	LUW1 G		6.10			25.00	27.00	87.00			1.50	

Table 26 – LUW2 - total soil analysis.

	Sample	Moisture Content (%)	Bulk density (g/cm³)	pH H2O	Clay (%)	Fine silt (%)	Coarse silt (%)	Total silt (%)	Fine sand (%)	Coarse sand (%)	Total sand (%)	Crusting index	CEC (cmol (+) /kg)
0 - 30 cm	LUW2 A	34.66	1.52	7.32	10.41	18.90	52.44	71.34	13.90	4.35	18.25	2.43	6.80
	LUW2 B	33.18	1.59	7.02	11.18	18.51	53.21	71.71	12.62	4.49	17.10	2.42	8.00
	LUW2 C	32.90	1.48	7.24	10.86	18.23	53.15	71.38	13.14	4.61	17.75	2.35	7.36
	LUW2 D	32.76	1.45	7.22	10.42	17.37	52.87	70.24	15.04	4.31	19.34	1.68	7.55
	LUW2 E	33.20	1.48	7.21	10.52	17.92	52.20	70.12	14.40	4.97	19.37	2.13	7.89
	LUW2 F	32.41	1.55	7.42	10.79	18.88	51.25	70.13	13.27	5.81	19.08	2.33	7.88
	LUW2 G			7.60	10.50	19.00	51.60	70.60	14.90	4.00	18.90	2.07	
30 - 60 cm	LUW2 A	33.39	1.69	7.29	12.05	19.83	51.72	71.55	12.60	3.80	16.39	2.63	7.54
	LUW2 B	32.18	1.73	7.39	10.79	20.05	52.43	72.47	12.71	4.02	16.73	2.84	7.72
	LUW2 C	30.79	1.56	7.10	12.59	21.36	51.11	72.47	11.00	3.94	14.94	2.43	8.06
	LUW2 D	33.62	1.58	7.20	10.14	17.55	51.86	69.40	15.63	4.83	20.46	2.16	7.40
	LUW2 E	34.59	1.54	7.25	12.29	20.74	49.55	70.30	12.62	4.79	17.41	2.67	7.18
	LUW2 F	35.04	1.60	7.41	12.84	20.24	52.15	72.38	10.82	3.95	14.77	3.24	6.63
	LUW2 G			7.70	19.70	18.90	49.20	68.10	10.40	1.80	12.20	2.26	

	Sample	SOC (%)	pH KCl	Total N (%)	Fe (mg/100g)	K (mg/100g)	Mg (mg/100g)	Ca (mg/100g)	Mn (mg/100g)	Na (mg/100g)	P (mg/100g)	Al (mg/100g)
0 - 30 cm	LUW2 A	0.92	6.10	0.09	53.90	15.80	28.00	132.00	12.55	< 1.92	18.50	21.06
	LUW2 B	0.97	6.32	0.09	50.20	20.90	28.00	141.00	15.51	< 1.92	17.90	23.60
	LUW2 C	0.94	6.25	0.09	55.00	17.60	26.00	129.00	13.04	< 1.92	18.10	22.78
	LUW2 D	1.16	6.03	0.10	61.60	33.00	30.00	124.00	12.08	< 1.92	23.00	23.07
	LUW2 E	0.97	6.18	0.09	53.00	20.10	29.00	117.00	10.07	< 1.92	18.40	22.74
	LUW2 F	0.92	6.37	0.09	61.10	17.10	32.00	146.00	13.11	< 1.92	20.60	23.14
	LUW2 G		6.80	0.09		26.00	24.00	151.00			13.50	
30 - 60 cm	LUW2 A	0.88	6.08	0.07	42.80	16.20	24.00	124.00	8.90	< 1.92	11.70	21.62
	LUW2 B	0.69	6.31	0.07	39.20	16.70	23.00	127.00	10.76	< 1.92	13.10	20.31
	LUW2 C	0.69	6.14	0.07	49.80	16.30	23.00	133.00	10.24	< 1.92	13.50	24.95
	LUW2 D	0.99	6.10	0.09	60.40	35.00	24.00	110.00	11.33	< 1.92	23.00	24.53
	LUW2 E	0.72	6.09	0.07	43.70	21.60	28.00	109.00	7.49	< 1.92	11.00	25.65
	LUW2 F	0.36	6.08	0.05	28.50	14.60	23.00	124.00	2.30	< 1.92	2.50	27.06
	LUW2 G		6.40	0.05		19.00	25.00	175.00			1.10	

Table 27 – LUW3 - total soil analysis.

	Sample	Moisture Content (%)	Bulk density (g/cm³)	pH H2O	Clay (%)	Fine silt (%)	Coarse silt (%)	Total silt (%)	Fine sand (%)	Coarse sand (%)	Total sand (%)	Crusting index	CEC (cmol (+) /kg)
	LUW3 A	41.64	1.40	7.72	19.82	24.78	48.79	73.56	6.04	0.58	6.62	1.70	11.26
	LUW3 B	35.68	1.51	7.81	14.10	19.58	55.21	74.78	9.76	1.36	11.12	1.72	9.29
0 - 30 cm	LUW3 C	37.94	1.43	7.59	13.78	20.47	56.29	76.76	8.38	1.09	9.46	1.63	9.06
	LUW3 D	38.17	1.49	7.89	17.62	21.54	52.09	73.63	7.55	1.19	8.74	1.49	11.08
	LUW3 E	36.50	1.60	7.88	11.85	20.14	58.05	78.19	7.57	2.40	9.97	2.01	9.80
	LUW3 F	35.87	1.58	7.49	12.10	17.95	59.33	77.28	8.74	1.87	10.62	1.93	9.91
	LUW3 G			7.60	18.60	21.40	49.70	71.10	9.40	0.90	10.30	1.37	
	LUW3 A	41.46	1.40	7.50	19.43	24.00	50.29	74.29	6.16	0.12	6.28	2.89	10.92
	LUW3 B	36.71	1.54	7.72	13.57	19.39	55.85	75.24	9.82	1.37	11.19	2.17	8.59
30 - 60 cm	LUW3 C	37.03	1.54	7.48	14.63	20.79	54.68	75.48	8.73	1.16	9.89	2.92	7.57
	LUW3 D	37.74	1.56	7.84	15.29	27.14	50.08	77.23	6.67	0.81	7.48	2.42	10.91
	LUW3 E	36.37	1.56	7.87	13.26	21.84	58.12	79.97	5.15	1.62	6.77	2.82	9.15
	LUW3 F	34.65	1.54	7.69	14.13	17.27	60.43	77.70	6.84	1.33	8.18	2.81	10.77
	LUW3 G			7.80	20.40	21.60	48.80	70.30	8.80	0.40	9.20	2.38	

	Sample	SOC (%)	pH KCl	Total N (%)	Fe (mg/100g)	K (mg/100g)	Mg (mg/100g)	Ca (mg/100g)	Mn (mg/100g)	Na (mg/100g)	P (mg/100g)	Al (mg/100g)
	LUW3 A	1.04	6.56	0.12	52.80	27.00	22.00	240.00	11.54	< 1.92	7.80	24.73
	LUW3 B	1.21	6.91	0.13	77.40	29.00	16.70	290.00	17.97	< 1.92	27.00	17.81
0 - 30 cm	LUW3 C	1.26	6.63	0.14	52.60	29.00	15.70	220.00	11.03	< 1.92	11.20	19.59
	LUW3 D	1.34	6.91	0.15	60.70	35.00	19.00	320.00	17.61	< 1.92	23.00	20.45
	LUW3 E	1.03	6.85	0.11	59.00	19.00	15.30	260.00	15.98	< 1.92	17.80	18.40
	LUW3 F	1.10	6.66	0.12	56.20	22.00	14.50	250.00	18.59	< 1.92	16.70	20.84
	LUW3 G		6.80	0.15		32.00	19.00	295.00			12.80	
	LUW3 A	0.39	6.16	0.06	24.30	14.60	21.00	205.00	4.14	< 1.92	1.50	30.19
	LUW3 B	0.66	6.73	0.08	55.50	15.20	14.10	195.00	9.43	< 1.92	14.70	17.06
30 - 60 cm	LUW3 C	0.59	6.44	0.07	39.90	16.70	15.80	152.00	6.51	< 1.92	4.50	17.93
	LUW3 D	0.58	6.62	0.07	37.60	18.50	18.10	230.00	8.73	< 1.92	8.40	27.79
	LUW3 E	0.49	6.82	0.06	44.30	10.10	14.50	230.00	11.99	< 1.92	8.10	22.51
	LUW3 F	0.35	6.34	0.05	26.50	12.20	14.30	220.00	6.41	< 1.92	4.80	24.37
	LUW3 G		6.60	0.05		17.00	18.00	222.00			1.20	

Table 28 – LUW4 - total soil analysis.

	Sample	Moisture Content (%)	Bulk density (g/cm³)	pH H2O	Clay (%)	Fine silt (%)	Coarse silt (%)	Total silt (%)	Fine sand (%)	Coarse sand (%)	Total sand (%)	Crusting index	CEC (cmol (+) /kg)
0 - 30 cm	LUW4 A	24.81	1.67	6.18	9.11	9.49	25.44	34.93	51.56	4.40	55.96	1.23	6.05
	LUW4 B	33.41	1.60	6.54	9.17	8.79	20.24	29.03	57.52	4.28	61.80	0.90	6.91
	LUW4 C	26.63	1.74	6.15	9.06	9.82	21.52	31.34	55.63	3.97	59.60	1.19	6.16
	LUW4 D	33.43	1.56	7.71	9.53	8.77	19.83	28.60	57.96	3.90	61.87	0.86	8.28
	LUW4 E	28.60	1.60	6.12	9.42	9.42	21.86	31.28	55.57	3.72	59.29	1.11	6.31
	LUW4 F	29.63	1.64	7.04	9.05	8.68	21.13	29.80	56.98	4.16	61.14	1.04	6.80
30 - 60 cm	LUW4 A	28.58	1.79	6.49	7.21	7.59	15.94	23.53	64.59	4.66	69.25	1.51	5.39
	LUW4 B	31.35	1.69	7.45	9.78	8.28	16.94	25.21	60.85	4.15	65.00	1.11	7.24
	LUW4 C	28.43	1.71	6.56	9.09	13.63	24.99	38.62	47.84	4.45	52.29	2.34	6.43
	LUW4 D	28.21	1.66	7.89	8.59	9.34	20.55	29.89	57.27	4.24	61.52	1.24	6.85
	LUW4 E	30.65	1.72	6.84	8.77	5.48	9.50	14.99	72.47	3.77	76.24	1.10	6.23
	LUW4 F	27.00	1.45	7.44	10.51	9.38	20.65	30.03	55.44	4.02	59.46	1.25	6.68

	Sample	SOC (%)	pH KCl	Total N (%)	Fe (mg/100g)	K (mg/100g)	Mg (mg/100g)	Ca (mg/100g)	Mn (mg/100g)	Na (mg/100g)	P (mg/100g)	Al (mg/100g)
0 - 30 cm	LUW4 A	0.76	4.66	0.08	94.30	10.10	12.90	89.00	6.28	2.00	29.00	40.70
	LUW4 B	0.95	5.23	0.10	152.30	8.40	12.40	136.00	9.11	2.20	52.00	25.35
	LUW4 C	0.84	4.61	0.08	87.80	7.50	11.70	88.00	6.03	< 1.92	30.00	25.17
	LUW4 D	1.04	6.34	0.10	167.80	18.10	17.10	220.00	17.05	3.40	62.00	14.94
	LUW4 E	0.79	4.65	0.08	115.50	13.70	10.50	84.00	5.50	< 1.92	32.00	25.35
	LUW4 F	0.87	5.60	0.09	101.60	6.30	9.40	147.00	9.49	< 1.92	38.00	16.77
30 - 60 cm	LUW4 A	0.40	5.09	0.04	49.10	8.20	11.40	90.00	1.10	< 1.92	3.10	51.05
	LUW4 B	0.42	6.06	0.05	62.50	15.60	17.00	147.00	6.31	3.70	20.50	24.06
	LUW4 C	0.43	4.95	0.05	53.50	9.70	14.50	100.00	3.97	< 1.92	8.40	26.22
	LUW4 D	0.74	6.86	0.07	136.40	18.50	16.10	240.00	10.02	5.80	48.00	17.33
	LUW4 E	0.40	5.22	0.05	80.90	14.90	11.90	96.00	3.20	< 1.92	18.90	21.98
	LUW4 F	0.56	6.11	0.06	112.30	7.10	13.60	150.00	5.44	< 1.92	20.50	14.06