-

Use of data in sustainability analysis

Assessing potential of geospatial data in peat field identification

Lappeenranta–Lahti University of Technology LUT

Software Engineering and Digital Transformation

PROPHET Double Degree – Sustainable ICT Solutions of Tomorrow

2024

Nelli Kemi

Examiners: Professor Jari Porras (Lappeenranta–Lahti University of Technology LUT)

Associate professor Lene Tolstrup (Aalborg University, AAU)

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Engineering Science

Software Engineering

Nelli Kemi

Use of data in sustainability analysis: Assessing potential of geospatial data in peat field identification

Master's thesis

2024

66 pages, 16 figures, 10 tables and 1 appendix

Examiners: Professor Jari Porras (LUT) and associate professor Lene Tolstrup (AAU)

Keywords: Geospatial data, peat fields, data science, sustainability analysis

In Finland, drained peatlands are important for the agricultural sector, but they are also important sources of greenhouse gases. Therefore, they are targeted by nature restoration regulations. To better direct these regulations and incentives, a deeper understanding of peat fields is needed, especially a way to determine peat fields belonging to real estate.

One potential solution could be the integration of geospatial data. Combining geospatial data from multiple sources could provide a way to determine the locations of peat fields and provide insight into their state. Similar combinations have yet to be widely researched.

This thesis aimed to investigate the topic by collecting open geospatial data from different Finnish institutes and spatially combining the datasets to collect needed information. Then, a solution was developed for displaying the results in a way to best allow for further sustainability analysis to be conducted. The resulting solution's sustainability impacts were evaluated using the Sustainability Awareness Framework (SusAF).

A definition for peat field parcels was established as well as an overview of their distribution and the crops cultivated on them. The proposed application was estimated to have a variety of potential impacts on the different dimensions of sustainability in the short term and the long term. The most important effects included the possible evaluation of real estates with peat fields and the possibility to direct decision-making.

Further research could be conducted on the accuracy and applicability of the proposed peat field definition as well as their economic importance.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT Teknis-luonnontieteellinen

Tietotekniikka

Nelli Kemi

Datan hyödyntäminen kestävyysanalyysissä: Paikkatiedon hyödynnettävyyden arviointi turvepeltojen tunnistamisessa

Tietotekniikan diplomityö

2024

66 sivua, 16 kuvaa, 10 taulukkoa ja 1 liite

Tarkastajat: Professori Jari Porras (LUT) ja avustava professori Lene Tolstrup (AAU)

Avainsanat: Paikkatieto, turvepelto, datatiede, kestävyysanalyysi

Ojitetut turvemaat ovat Suomessa tärkeitä monille maanviljelijöille, mutta ne ovat myös merkittäviä kasvihuonekaasupäästöjen lähteitä sekä ennallistamisasetusten kohteina. Jotta tarvittavat toimenpiteet ja mahdolliset avustukset voidaan ohjata oikeisiin kohteisiin, tarvitaan lisäymmärrystä turvepeltoihin liittyen, erityisesti keino yhdistää turvepellot kiinteistöille.

Yksi mahdollisuus ymmärryksen lisäämiseen voisi olla paikkatiedon hyödyntäminen. Paikkatiedon yhdistäminen useista lähteistä voisi mahdollistaa turvepeltojen sijaintien määrittämisen sekä tarjota lisätietoa niiden tilasta. Tällaisia yhdistelmiä ei vielä ole juurikaan tutkittu.

Tämän työn tavoitteena oli tutkia aihetta keräämällä usean suomalaisen instituutin tuottamaa paikkatietoainestoa ja yhdistämällä näitä tarvittavan tiedon keräämiseksi. Tuloksien visualisointia varten kehitettiin sovellus, jolla pyrittiin mahdollistamaan tulosten hyödyntäminen erilaisissa kestävyysanalyyseissä. Kehitetyn sovelluksen kestävyysvaikutuksia arvioitiin Sustainability Awareness Frameworkin eli SusAF-kehyksen avulla.

Työssä kehitettiin määritelmä turvepeltolohkoille, jonka avulla selvitettiin niiden osuus Suomessa sekä niillä kasvatettavat kasvilajit. Kehitetyllä sovelluksella arvioitiin olevan useita mahdollisia vaikutuksia kestävyyden eri pilareilla lyhyellä ja pitkällä aikavälillä. Merkityksellisimpiä vaikutuksia olivat mahdollinen turvepeltoja sisältävien kiinteistöjen tunnistaminen sekä mahdollisuudet vaikuttaa päätöksentekoon esimerkiksi tukiin liittyen.

Jatkotutkimusta voitaisiin tehdä erityisesti turvepeltolohkomääritelmän oikeellisuuden ja laajemman soveltuvuuden arvioimiseksi sekä turvepeltojen taloudellisten vaikutusten suhteen.

ABBREVIATIONS

CH_4	Methane
CO_2	Carbon dioxide
CRSD	Corporate Sustainability Reporting Directive
EU	European Union
GDPR	General Data Protection Regulation
GHG	Greenhouse gas
GIS	Geographic Information System
GTK	Geological Survey of Finland
Luke	Natural Resources Institute Finland
MoSCoW	MoSCoW prioritization method
N_2O	Nitrous oxide
NLS	National Land Survey of Finland
PDF	Portable Document Format
SusAD	Sustainability Analysis Diagram
SusAF	Sustainability Analysis Framework
UNEP	UN Environment Programme

Table of contents

Appendices

Appendix 1. Most and least common crops grown on Finnish peat fields

1 Introduction

Peatlands might only cover a few percentages of land area worldwide, but they are very important ecosystems containing up to a third of global soil carbon [1]. Maintaining carbon in the soil is essential for controlling climate warming. However, peatlands can be drained, and the soil carbon exposed for various reasons, such as for agricultural purposes.

Organic soils used for agricultural purposes are a source of environmental challenges due to their potential to emit higher levels of greenhouse gases (GHG) compared to inorganic lands. The organic material, peat, that they consist of is formed when acidic conditions and lack of microbial activity prevents plant materials from fully decaying [2]. Due to this lack of decay, peat lands' soil contains nitrogen and substantial amounts of carbon, both of which are potential sources for known greenhouse gases carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) [2]. Cultivation on drained peatlands is responsible for 32% of greenhouse gas emissions of all croplands although crops from these peat fields cover only 1.1% of calories from crops consumed by humans [3].

Past research has found that warmer winters can potentially lead to increased N_2O emissions from peatlands, with up to 106% more N_2O emissions during a warmer winter compared to a winter with temperatures closer to the average [4]. They found CO_2 emissions to also be higher during the warmer winter although not as significantly as N_2O. These findings suggest that climate change has the potential of creating a vicious circle: global warming results in more N_2O and CO_2 being emitted, in turn increasing the rate of climate change.

One key factor in mitigating the effects of peatlands could be a better understanding of their state and locations. This could be achieved with the help of technology. In recent years, information technology has developed fast leading to software solutions and data analysis tools, such as Geographic Information Systems (GIS), that can be used to discover information needed to improve sustainability.

GIS is a computer system that can be used to collect, store, analyse, and display geographic information on the Earth's surface [5]. It can be considered to tie knowledge to a grid through the combination of geolocation and different types of qualitative statistics [5]. The type of data GIS uses is known as geospatial data: information that describes features with a location on or near the Earth's surface by combining information related to the location and some attribute with temporal information [6].

Some research has been conducted on peatlands using geospatial data in the past. In 2018, PEATMAP, a global map combining a variety of sources into a single peatland map [7], was developed and more recently in 2022 the UN Environment Programme (UNEP) published a reported on the current global scale and distribution of peatlands [1]. However, there is still a lack of studies focusing on the peatlands used for agriculture specifically.

Although peatlands are not a main soil type for agriculture, rewetting these drained lands could lower the European Union's (EU) GHG emissions from agriculture by 25% [1]. This shows how understanding geo-environmental factors influencing agriculture in a spatial context is needed to improve sustainable agricultural practices as detailed, trustworthy, and up-to-date geospatial data can be used to determine the agricultural conditions required for sustainable agriculture [8].

By now many companies within and outside of the agricultural industry have started to implement sustainability strategies, such as strategies including actions to increase biodiversity, to improve sustainability. Some of these are based on voluntary information but regulations and directives often provide a baseline for which data needs to be collected and how this data can be used. The directives affecting different companies vary based on the location and sector of operation as well as the size of the company.

In Europe. the EU has implemented many directives and regulations to guide actions toward sustainability. For example, all Member States have committed to making Europe the first climate neutral continent by 2050, requiring at least a 55% reduction in emissions compared to levels in 1990 [9]. Achieving this goal requires adhering to multiple strategies such as the European Green Deal which aims to make the EU a fair and wealthy society [10].

The Green Deal is a set of strategies combining a variety of aspects including the environment, agriculture, sustainable financing, and more [10]. It includes strategies such as farm to fork, and the EU biodiversity strategy that are closely related to the agricultural sector. The farm to fork strategy aims to shift EU's food system toward a sustainable model by supporting sustainable food production through goals for organic farming and carbon capturing [11]. The strategy is closely linked to the biodiversity strategy which aims to preserve nature's biodiversity through goals such as restoration of degraded ecosystems by 2030 and annual funding directed at improving biodiversity [12].

Another goal of the Green Deal is for the EU to have a modern and competitive economy. Thus, the EU provides tools for economic sustainability, EU taxonomy being one of them. It is a part of EU's sustainable finance framework and a major tool in improving market transparency [13]. It helps companies and investors identify economic activities that are environmentally sustainable when aiming to make sustainable investments [14]. The taxonomy allows for companies both within the financial sector as well as outside of it to have a collective understanding of environmentally sustainable economic activities [13]. The EU Taxonomy is currently not a set of mandatory requirements, but it is expected to eventually encourage transitions towards sustainability [14].

One directive helping to evaluate companies' sustainability performance in relation to the green deal is the Corporate Sustainability Reporting Directive (CSRD) [15]. The goal of this directive is to increase transparency and sustainability within supply chains. The CSRD requires companies to enclose information on both risks and opportunities resulting from environmental and social issues they have identified as well as their impact on the environment and humans [15]. Reporting related to agriculture within this directive is especially relevant for companies within the agricultural sector but other companies having agriculture as a part of their supply chain might need to report on the fields' impacts as well.

One of the most important regulations regarding data usage within the EU is the General Data Protection Regulation (GDPR). It aims to protect natural persons regarding how their personal data is handled [16]. It affects all data processing and analysis when dealing with personal data in EU. Fields used for agricultural purposes are often owned by natural persons and when aiming to utilize data related to the lands, GDPR can limit the way the related data can be processed.

Most recently, in June 2024, the EU adopted a new regulation known as nature restoration law [17]. All member states have to restore 30% of drained peat fields by 2030 and rewet a quarter of them [18]. By 2050 half of peat fields must be rewetted although the member states are allowed to set themselves lower goals with valid reasoning [18]. For countries like Finland, where peat fields have been estimated to cover around 11% of the total agricultural area [19], this will be an important consideration. However, it has been suggested that removing some 20 000 hectares of peat fields from use in Finland would not have a negative impact on food production but would noticeably lower the country's total GHG emissions [20].

1.1 Motivation and objectives

As sustainability and data are both globally significant areas of interest, their connection becomes more and more important. Regulations and directives set a foundation for information that must be collected, analysed and reported. For companies aiming for a sustainable reputation, going above and beyond these regulations can be beneficial. Hence, these companies wish to invest in researching new topics and points-of-views related to the different aspects of sustainability.

This is also the case for an organization that proposed the directions for this study. In 2023 the Natural Resources Institute Finland (Luke), the National Land Survey of Finland (NLS) and the Geological Survey of Finland (GTK) conducted a study in collaboration with the University of Turku to define and identify peatland sites in Finland, resulting in a dataset including locations for the identified areas [21]. The applicability of the resulting dataset in identifying peat fields could be of valuable insight but has yet to be tested in further implementations. Hence, the organization is interested in finding out if the dataset can be used in combination with other open data to extract and visualize meaningful information regarding peat fields in Finland on a regional and a real estate level as so far the latter has not been possible. The results could then be used as a basis for further analysis related to, for example, the importance of the different types of agricultural soils. Additionally, they are interested in knowing if the results could be used together with their internal data.

The environment this research is conducted in is visually depicted below in Figure 1. The image provides a visual representation of the previously described regulatory landscape as well as the main concerns and interests resulting in this work.

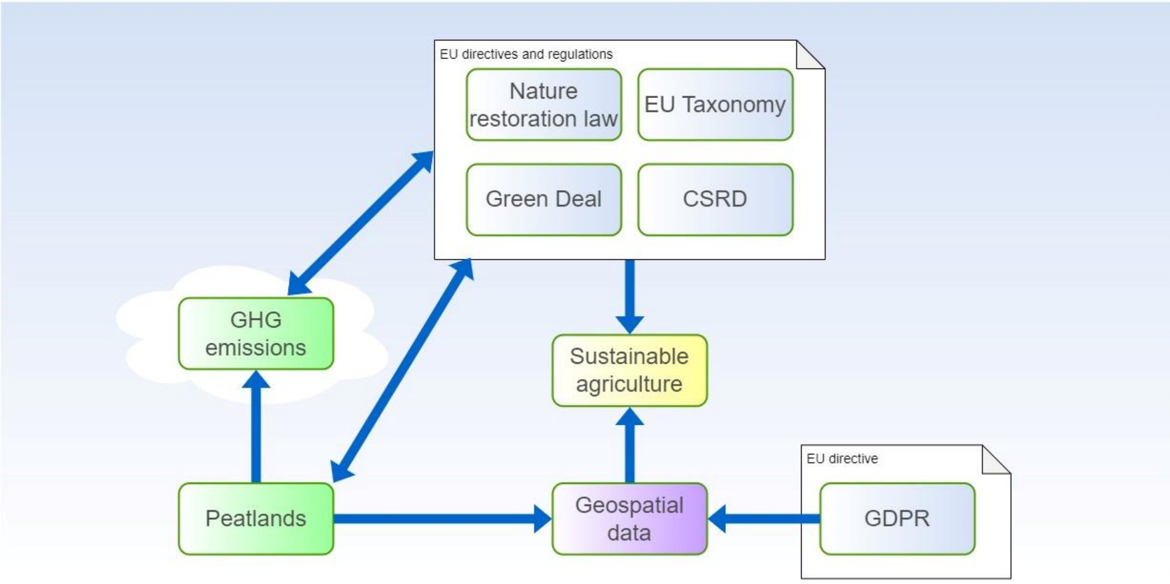


Figure 1. Regulatory environment for the research

The proportionally higher GHG emissions from peat fields has resulted in them being one of the targets of many directives. Finland has proportionally more peat fields than many other European countries [22], which creates challenges with the restoration law. Nevertheless, over half of the emissions from the Finnish agriculture are from peat fields [22] which is why action on these fields is important as also required by the nature restoration law. On the other hand, it might be challenging for farmers to stop using their peat fields for agricultural purposes as there might only be a limited possibility of other sources of income [22]. This is where identification of peat fields could help with determining a sustainable rate of rewetting and a fair distribution of financial support.

The objective of this study is to investigate the application of the new peatland dataset in combination with other open datasets provided by NLS and the Finnish Food Authority. This study is conducted in collaboration with the case organization with an interest in leveraging more detailed information of peat fields and the applicability of these datasets in particular in doing so. One key aspect of this study is to determine requirements for when field parcels are considered peat field parcels as the definitions used are currently not universal [21].

The results will include a definition for a peat field within the datasets, statistics related to the current state of peat fields in Finland, and a prototype of a visualization dashboard for displaying the results. The results will be stored in a way that they can be compared to and combined with the case organization’s internal data, but the results will also be usable independently of any additional data.

1.2 Research questions

The aim of this research is to investigate the usability of peatland data and the information that can be extracted from the data for sustainability analysis. Hence, two main research questions will be explored:

- How can geospatial data of peat fields be used to provide users with meaningful information for sustainability analysis?
- What are the benefits of using data in sustainability analysis?

These two questions complement one another by covering different aspects of the problem. These questions will be explored through a case study. The second question will require interviews to be conducted within the case organization.

To help determine an answer for the first question, a set of sub questions were defined. These questions will guide the data analysis and help better understand the gaps in current research and the importance of this research. These sub questions are:

1. What is the status of peat fields in Finland?
2. How has geospatial data been used in analysis related to peatlands?

To better understand the current state of peat data in sustainability analysis, an exploration through a literature review will take place. Based on the results, research gaps can be identified and a direction for data analysis defined. Following this set of directions data collection, data processing, and data analysis will be conducted to answer the main research questions, without omitting an analysis on the implications of these results. A more detailed description of the research methods will be provided in Chapter 3.

1.3 Structure and limitations

This work focuses on identifying the current state of peat fields specifically in Finland. Other areas or the applicability of the proposed methods in other areas will not be considered at this time. The parameters that can be used for the identification of peat fields are limited by the available datasets as no field data will be collected during the research process. The geospatial datasets are limited in resolution to 10x10m which sets limitations to the accuracy of the resulting data.

The extent of analysis and application is limited by the case organization's needs as well as the resources available for this research. Rather than aiming to investigate the different tools and frameworks available for data analysis the focus of this research is on data science: discovering and interpreting knowledge from data to be used as guidance for decision making or strategic planning [23].

The literature review conducted on related research will adapt the guidelines for a systematic literature review, however, as it is not the focus of this work, a lightweight approach will be taken for the process. Similarly, as the main technical focus of this work is on examining data, the software lifecycle related practices will be less rigorous than in a full software development process. The proposed solution will be on the Proof-of-Concept level and, hence, the user interface and user experience design within this research project will be minimal.

Next, this paper will continue with the analysis of related works in Chapter 2, after which the methods used for this research are described in Chapter 3. In Chapter 4 the results obtained within this research are presented. Chapter 5 provides insights into the results and their implications, and prospects for further studies before the work is concluded in Chapter 6.

2 Related works

In this chapter a closer look will be taken at related scientific works. This is done by adapting principles from the process of a systematic literature review in software engineering as defined by Kitchenham and Charters in [24]. According to them, the goal of such reviews is to answer questions of significant importance to practitioners as well as potentially change current practices. However, with thesis works the goal of the literature review is often different, the aim being to identify existing basis for the research project to be conducted [24], and this approach is taken within this work as well.

The systematic literature review process can be divided into three main steps: planning, conducting and reporting [24]. Kitchenham and Charters described the planning stage to cover the design of the review process including identifying the need and the protocol, which includes research questions and a search string. After this, the conducting step includes initial research to determine a search strategy, selection of the studies, and extraction and synthesis of data. Lastly, in the reporting step the findings are written down for presentation.

2.1 Review protocol

For this work, the planning started with identifying databases to be used for the search. In the end two databases were selected: Scopus and Web of Science. These databases were chosen because of the variety of disciplines the included articles cover. This was an important criterion as the topic is an interdisciplinary topic with possibilities to cover many types of research.

Different search terms were investigated for use during the initial searches. Initially broad terms such as “data” and “sustainability analysis” were used to discover a baseline but more exact terms were needed to receive more focused results. Hence, “geospatial data” was identified as one key search term, as this study intends to investigate the usability of geospatial data of peat fields. At first, the term was used in combination with “peat*” to include all terms related to peat and secondly with “organic soil”, but these searches reached very few articles. The terms "geospatial" and "data" were then separated to allow for different ways of expressing the idea and allowing for more articles to be found whilst mostly remaining in the field of geospatial data. Therefore, the final combination of these terms became "(peat* OR 'organic soil') AND geospatial AND data". The query was applied to the title, abstract, and keywords of the items.

Additional inclusion criteria were defined to limit the number of articles to match the available resources and to discover the most relevant works. Only research articles written in English and published in journals were included. Additionally, the whole article needed to be available for inspection. Similarly, only articles with a focus on peatlands and their spatial distribution were included. Geospatial data or geospatial analysis techniques needed to be mentioned in the methodology, studies using only soil measurements were excluded.

In total 16 articles matching the inclusion criteria were identified. These articles served as a start set for forward snowballing [25]. According to the methodology, more articles are discovered by identifying citations of the articles within the starting set, for example, using Google Scholar. The information provided by Google Scholar is then studied to determine whether the article should be included. As hundreds of citing articles were discovered, an additional search term "geospatial data" was applied to these articles. Otherwise, the same inclusion criteria as for the start set were applied. An additional number of 8 articles were identified this way. In the end, a total of 24 articles were closely examined for this section.

2.2 Data synthesis

The collection of studies was a diverse set examining organic soils from multiple directions and in many parts of the world. The studies focused on a variety of topics and had different goals. Common topics within the works were identified through a thematic analysis. An inductive approach was taken where the articles were analysed to determine the studies' topics and these topics were then grouped into common themes encompassing multiple topics [26]. The distribution of articles according to the identified themes can be seen below in Table 1. In this categorization one article can be included under more than just one theme.

Table 1. Themes of analysis

Theme	Quantity	References
Peatland areas disturbance	2	[27,28]
Distribution of peatlands	4	[7,29-31]
Peat layers	2	[29,32]
Crops	2	[31,33]
Land-cover changes	3	[34-36]
Ecosystems	4	[35-38]
Hydrology	2	[32,39]
Monitoring / risk register	4	[34,36,38-40]
Restoration / preservation	6	[30,35,36,38-40]
Carbon stocks	2	[41,42]
Climate change / emissions	6	[30-32,43-45]

The most common themes within the studies seemed to be peatlands’ impact on climate change and GHG emissions as well as restoration or preservation strategies related to them. Other common areas of interest were the distribution of peatlands and their ecosystems. Only one study looked at crop rotations and the impacts of the crops within peatlands [33] although a second study also highlighted the need for integration of crop information into geospatial datasets to improve GHG emission estimates from drained organic soils [31]. Very little attention was also paid to drained peatlands in use for agricultural purposes or to create a definition for peat fields.

Additionally, the outcomes of all studies were statistical. Most studies reported the results of the analysis whilst some studies had made the resulting datasets and maps available to the public, for example, [7]. A few studies, such as [28], also highlighted the usability of the results as tools in planning and decision-making. However, none of the studies aimed to develop or investigate possibilities to develop solutions for displaying the results of the analysis or building tools that could be used to do so.

Most studies outlined a smaller geographical area to conduct the research on. The distribution of the geographical study areas is shown below in Table 2.

Table 2. Spatial distribution of researched areas

Area	Quantity	References
Finland	4	[27,33,39,46]
Canada	3	[30,32,47]
USA	2	[37,45]
Russia	2	[42,43]
Ireland	4	[28,36,38,40]
Peru	1	[41]
Rest of Europe (Latvia, Germany)	2	[29,48]
Southeast Asia (Indonesia, Malasia, Thailand)	4	[34,35,44,49]
Global	2	[7,31]

Only two studies did not focus on a specific geographical area but rather aimed to cover the whole planet [7,31]. Out of the studies' areas, the most common for this group of studies were Finland, Ireland, and Southeast Asia. These study areas correlate well with the findings of [7] and the UNEP [1] regarding the global distributions of peatlands and highlights the importance of these lands for the researched areas.

As most studies focused on a limited geographical area, the second most common research method identified was case study. The most common research method was geospatial analysis as the inclusion of geospatial data or geospatial analysis was required by the inclusion criteria. The distribution of articles into the other most often appearing research methods is shown below in Table 3.

Table 3. Research methods

Research methods	Quantity	References
Case study	22	[27-30,32-49]
Field study	5	[32,46-49]
Statistical analysis	6	[28,31-33,38,41]
Temporal analysis	3	[33,44,45]

The two studies not focusing on a specific geographic area were conducted as a meta-analysis [7] and using a pre-developed statistical methodology [31]. All other studies were included in the case study category. Limiting the study area helps with controlling the size and number of datasets and can be a useful strategy for conducting these types of studies. However, the application of the achieved results in wider areas or on a global scale is underrepresented. Additionally, the two studies that did not limit their study area were also able to achieve meaningful results, suggesting that the limitation is not required in all situations.

Out of the studies taking the case study approach, six included principles of field studies. In these studies, the researchers had collected their own samples from nature. In one study peat layer thickness was measured by manually extracting soil cores and probing the soil using a metal rod, and the resulting data was used as a part of training data for a machine learning model [47]. Metal rod probing was also used to measure peat layer thickness in [46] whereas [49] sampled the soil to determine moisture levels for accuracy verification. Samples from rewetted peatlands were collected in [48] to analyse the soil properties, and burn severities were measured in [32] as a part of their data collection.

The studies conducting statistical analysis had different approaches, one study using multiple statistical modelling methods [43] while others used, for example, chi-square also known as goodness-to-fit tests [28,32,41]. Other statistical approaches included using a new statistical framework [31], and tools such as SAS software [48]. On the temporal side studies compared historical and current data to analyse changes having taken place over time.

Similar to the research methods, a variety of types of data can be utilized when conducting geospatial analysis. The types of data used within the articles discovered are presented below in Table 4. All articles included more than a single type of data.

Table 4. Data types

Research methods	Quantity	References
Remote sensing data	13	[7,27,28,32,34,35,37,39-41,46,47,49]
Geospatial data	6	[7,32,33,38,42,43]
Field data	6	[32,41,46-49]
Statistical data	8	[29,31,34,36,37,42,44,45]

In the table, remote sensing data and geospatial data are classified separately. This distinction was made as remote sensing can be used to create geospatial data. Hence, remote sensing data includes data such as satellite imagery [27] whereas geospatial data includes spatial data with descriptive attributes, for example, the digital soil maps used in [7]. Similarly, field data included data collected by the researchers as previously described but also data collected by previous research like the data on ecosystem types used in [41]. Statistical data refers to other data not identified as geospatial data based on the description, such as emergency department visitation data [37].

Some studies utilized data from many sources, for example, [32], whereas others were focused on field data like [48]. The combination of data from different sources highlights the importance of both data quality and quantity. In order to come to conclusions, enough data needs to be available. Similarly, the quality of the data is important for the meaningfulness of the results. Both aspects can be better ensured when data is retrieved from multiple sources, as was the case with these studies. Additionally, including different types of data allows to add depth to the analysis. Nonetheless, certain conclusions could also be drawn with fewer types of data.

Overall, there is strong evidence that case studies and data analysis using data from multiple sources is a feasible method for topics covering geospatial aspects of peatlands. Although Finland was one of the most common study areas within this collection of articles, little attention had been paid to agricultural lands specifically. Similarly, no studies were identified on the distribution of peat fields among farmers or development of an application for displaying the obtained results. Hence, there is a gap in connecting these pieces of information.

3 Research methods

This chapter will describe the steps taken to answer the research questions introduced in Section 1.2. The project methodology is first described on a high level before practical steps towards the technical solutions are presented. A visual representation of the research methods and the research flow is shown below in Figure 2.

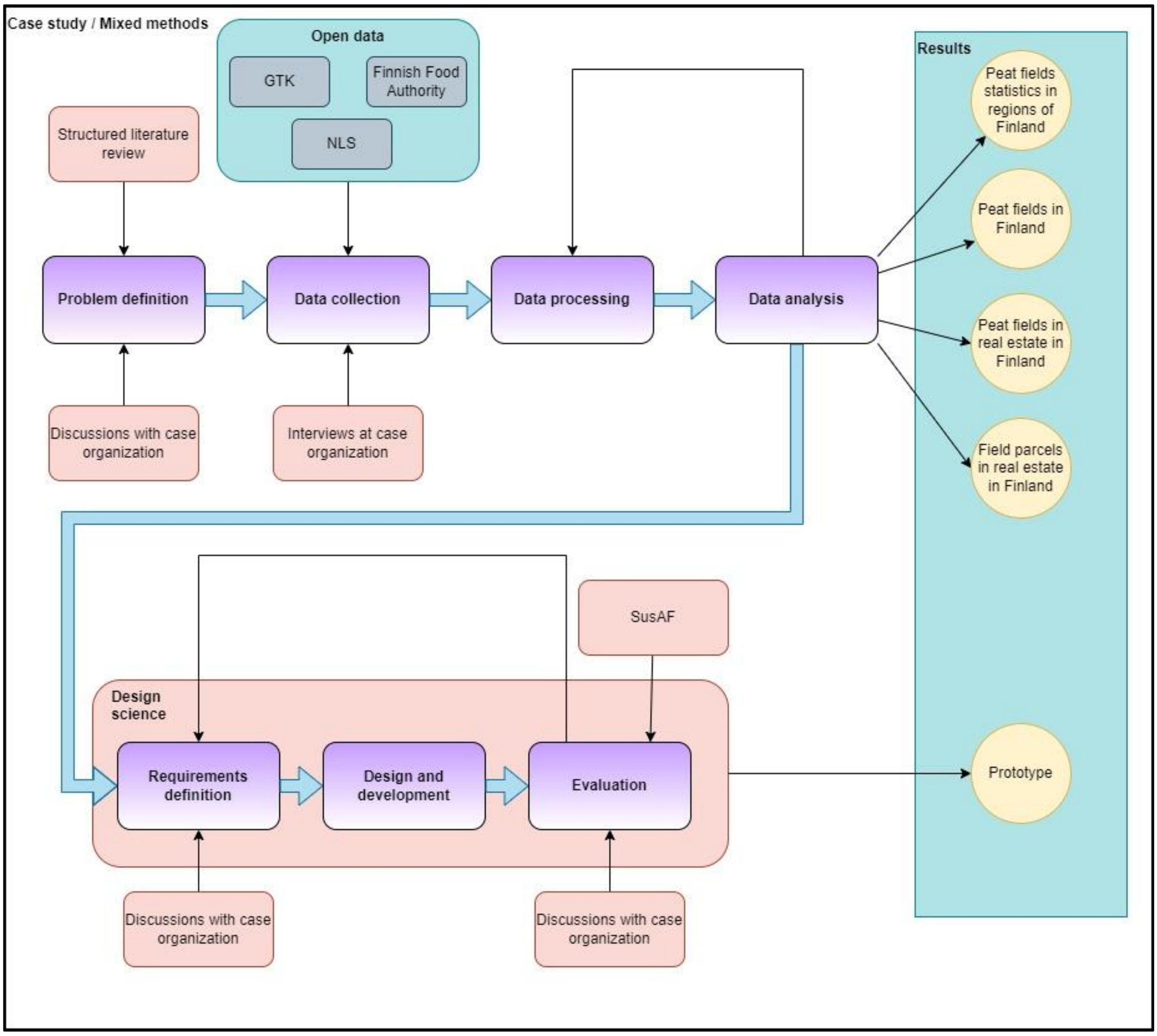


Figure 2. Visual presentation of the research process

The main steps within this research include problem definition, data collection, data processing, data analysis, requirements definition, design and development, and evaluation.

Although a fairly linear process will take place, iterative approach will be taken during data processing and analysis to ensure everything needed is available in the resulting datasets. Similarly, the requirements definition, design and development, and evaluation will be done iteratively to ensure that the developed prototype will be approved by the case organization.

The main methods this thesis follows are mixed methods and case study. Mixed methods combines aspects of both qualitative and quantitative research to provide more depth than a single method would [50]. On the qualitative side, the study will be conducted as a case study in collaboration with the case organization and using peat fields a specific case for sustainability analysis. Similarly, the state of sustainability analysis at the case organization will be explored and these results will be described in a qualitative manner. The datasets containing the final data also include qualitative information through the attributes assigned to the geolocations, such as information about the crop types.

This study also follows the principles of design science with the aim to create an artefact that can be used by people to solve practical problems of general interest [51]. According to this definition, design science aims to not only produce a novel artifact but also knowledge about its impact on the environment. This will be the main method used to solve one of the main research questions of this project: “How can geospatial data of peat fields be used to provide users with meaningful information for sustainability analysis?”

The principles of design science include formulating a problem statement, determining stakeholder goals and requirements, and evaluating proposed artefacts [51]. The problem statement for this research was introduced in Chapter 1 whereas stakeholder goals and requirements will be determined next. To help answer the second main research question, “What are the benefits of using data in sustainability analysis?”, a round of interviews at the case organization will be conducted. These interviews will provide a baseline for determining the stakeholder goals. The results of the interviews will also contribute to the qualitative aspects of this study through written summaries of the topics discussed.

Based on the results of the interviews as well as additional discussions with key stakeholders at the case organization, the need for this research can be identified and requirements for the artefact to be built can be defined. The process including but not limited to defining the system requirements is also known as requirements engineering. Out of the generally accepted parts of requirements engineering this study will incorporate collection of requirements from stakeholders, compiling and establishing the system requirements, and reporting the requirements [52]. As this research only aims to develop a proof-of-concept level prototype, requirements tracing, tracking and management throughout the software lifecycle are not considered central in this case.

Prioritization of the identified requirements will be done using the MoSCoW method. It classifies requirements in the order of importance from the most important to the least important: “Must” for requirements that must be fulfilled, “Should” for requirements with a high priority, “Could” for requirements that would be preferred, and “Won’t” for requirements that can be postponed [53].

Based on the defined requirements, designing and implementing the application will take place. However, to build a visualization dashboard, the data to be visualized must first be analysed. This analysis will provide an answer to the research question “What is the status of peat fields in Finland?”. In order to answer the question, data will be collected from multiple sources before being processed and combined in a meaningful manner to provide statistics on the distribution of peat fields in Finland and in Finnish real estate as well as information regarding the crops grown on the fields.

Once the development of the solution is finished, Sustainability Analysis Framework (SusaF) will be used to analyse and evaluate its impacts on the different aspects of sustainability. SusAF was created as a tool for sustainability design, and a workbook has been created to help utilize the framework in eliciting and analysing potential sustainability effects of IT solutions [54]. This thesis will follow the methodology from the sixth version of the workbook [54] which includes the following four steps: warm-up, capture, analysis and synthesis.

- The goal of the warm-up is to get familiar with the framework and the solution being analysed including possible already known sustainability effects. As this analysis will not include a team or any pre-conducted analysis, the warm-up is not central in this case.

- The capture step is the most important one for this study. In this step, all participants are expected to brainstorm new potential effects which are then classified according to their likelihood and impact. Five different sustainability dimensions are considered: social, individual, environmental, economic, and technical. Questions are provided for all dimensions to guide the brainstorming session.

- The results of the capture step will be used during the third step, analysis. During this step the orders of the effects are defined, and chains of effects are identified. These can then be displayed on a Sustainability Awareness Diagram (SusAD).
- Lastly, in the synthesis step the effects are translated into opportunities and threats. Actions to be taken to mitigate the threats and leverage the opportunities are identified. In this thesis, this step will be covered during the discussion chapter while discussing the results and possibilities for further research.

3.1 Data collection

The data collection phase encompassed two distinct steps. First, information was gathered about the practices at the case organization through interviews. Then, the geospatial data needed for the analysis was collected.

3.1.1 Interviews

To better understand the current practices and possibilities for improvement regarding sustainability analysis at the case organization, two employees were interviewed. Both employees were involved in sustainability analysis and reporting, one of them having worked more directly with using data for said projects and the other in a leading position.

The interviews were conducted on and recorded and transcribed by Teams. The interview consisted of open-ended questions that covered three themes: the interviewees background in sustainability analysis at the case organization, the current practices regarding sustainability analysis at the case organization, and directives relevant to the case organization. The interviewees were encouraged to share any experience and knowledge they found relevant to the topics.

It was established that the company has mainly focused on reporting and analysis required by directives and regulations such as the EU taxonomy and more recently the CSRD. Most of the reporting has been done to fulfil external demands whereas analysis has been used to guide internal operations. Certain demands for analysis have also come from other external sources such as the market and society, as the public has been becoming more aware of sustainability.

Regarding the data used for the analysis, mostly open data has been used but some internally collected data is applied as well occasionally. To the knowledge of the interviewees, no previous analysis had been done around farms or agricultural fields and their sustainability. However, open data from the Finnish Food Authority had already been used for other purposes. This identified gap in analysis provides room for investigation of agricultural areas and especially peat fields and their connection to different aspects of sustainability especially considering the organization’s previous experience with some of the same and similar datasets.

Although no previous sustainability analysis had been conducted on farms and fields, certain environmental aspects had been of interest to the organization in the past. Analysis related to CO_2 emissions and risks related to the climate and nature were some noticeable topics covered. Similarly, social aspects such as diversity and equity of workforce had been monitored.

Although to this day most demands for sustainability analysis have come from outside forces, the organization has been able to use these analyses for their own purposes as well. In the same way, as the company has desires to know more about the location and quality of peat fields in Finland, an analysis related to these would be of use for internal purposes. With an interest in expanding their capabilities in sustainability analysis and the assumption that there is no research available on the usability of the field location data published by Luke, NLS, and GTK, the organization would be interested in finding out the quality and usability of the data for their use cases.

3.1.2 Open data

For this project, four primary types of geospatial datasets were used. One dataset was separated into three parts and one dataset had four versions available, each corresponding to a different year. Therefore, in total nine different datasets were used. All datasets were downloaded directly from their providers. The datasets are introduced below in Table 5.

Table 5. Data sources

Name	Description	Type	Source
Peatland site types of Finland 1.0/2023	Datasets including geospatial data of peatlands in Finland and the data points' accuracy ratings.	Tiff / raster	GTK Hakku
Agricultural parcel containing spatial data 2020-2023	Geospatial dataset including information about the agricultural field parcels in Finland.	GeoPackage / vector	Finnish Food Authority
Cadastral index map	Dataset including information about real estate in Finland.	GeoPackage / vector	NLS
Division into administrative areas	Dataset including information about the regions of Finland.	GeoPackage / vector	NLS

Peatland site types of Finland 1.0/2023 is the new dataset by Luke, GTK, and NSL. It included three separate files, one containing information about drained peatlands, the second undrained peatlands, and the third accuracy ratings for the data. All datasets included pixel values each corresponding to a 10x10m area. The drained and undrained peatlands included values corresponding to the classification of the soil type. The accuracy rating included a value corresponding to an accuracy rating value that was accompanied by a PDF file containing accuracy matrixes expanding on the accuracy for the associated pixel points.

The agricultural parcel containing spatial data included field parcels for each year. The information contained for each parcel varied slightly between the different years, but the parameters important for this research were present each year: parcel identifier, crop grown, field area, and parcel geometry.

The cadastral index map included information about Finnish real estate, most importantly their identifiers and geometries. Additionally, the dataset included the type of real estate that could later be used to filter out real estate that are not wanted in the final results.

Lastly, the division into administrative areas included multiple layers of data. One of the layers included regional level information which was used in this research. The dataset included the regions' names in Finnish and their geometries.

3.2 Data processing and analysis

The data processing and analysis included multiple steps. First, the data was processed in a way to allow for calculation of quantitative statistics. This process started with pre-processing the datasets to a usable format, continuing with defining metrics for peat field identification, and ending by connecting the identified peat fields to the regions of Finland and Finnish real estate. Then qualitative attributes were investigated. All resulting datasets were stored in individual Apache Parquet files.

3.2.1 Peat field identification

To get the peatland sites datasets into usable format, QGIS was used for the processing. The Tiff files were uploaded to QGIS for filtering and transformation. The other datasets did not require additional processing to get them into a usable form.

First, the interesting data points within the peatland datasets needed to be identified. Visualization of one of the three datasets included in the peatland datasets, locations of undrained lands, is shown below in Figure 3.

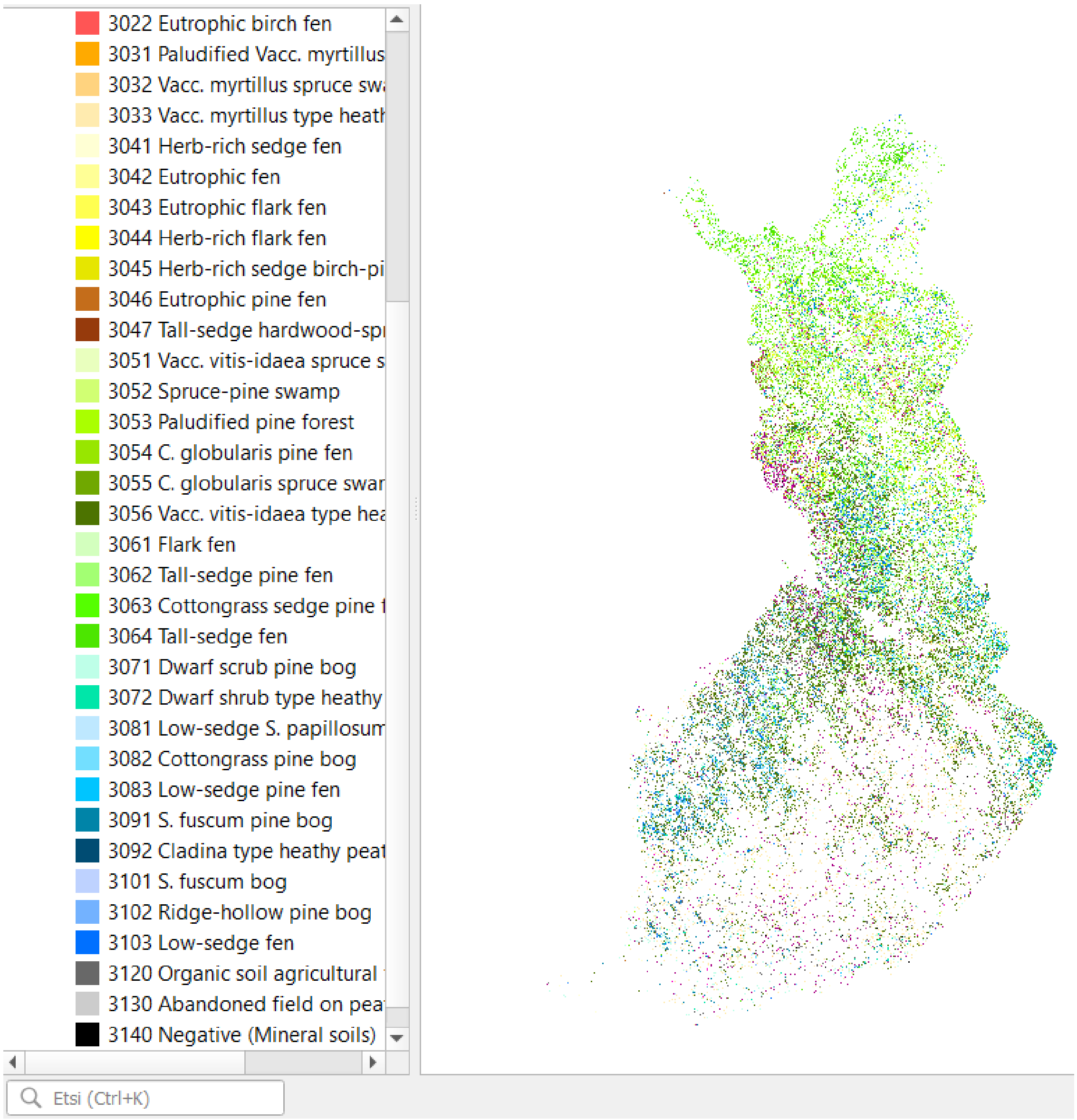


Figure 3. Drained peatlands

On the left is shown a snippet of the list of available pixel values with the name of the peatland type. The interesting pixels for this research are the pixels with the value 3120: Organic soil agricultural fields. In the other dataset including undrained lands the corresponding value was 2120.

The raw data within these files consisted of x and y coordinates for each pixel as well as the value corresponding to the type of peatland. As the aim is to identify the areas corresponding to peat fields, the pixels needed to be grouped into unified areas known as polygons. This was done by first filtering out other values besides 3120 and 2120 in the two datasets in

QGIS. This way two new datasets were obtained, each containing values 0 and 1 for each pixel: 1 if the original value was the desired one (3120 or 2120 depending on the dataset in question) and 0 for the others. Then, a vectorization transformation was done in QGIS to combine the individual pixels into polygons based on the values 0 and 1. This vectorization resulted in some incorrect geometries as certain polygons were self-intersecting which were still fixed in QGIS using the "Fix incorrect geometries" tool. The vectorization process was also done to the accuracy ratings in QGIS to obtain polygons with the corresponding accuracy rating values as attributes. It was then possible to read these files in Databricks and continue processing the datasets there using Python mostly with Pandas and GeoPandas.

When starting the process of identifying peat fields, it was important to determine which points and polygons from the data to include and which to exclude in the analysis. This was done with the help of the accuracy matrixes provided with the dataset as well as through investigation of the range of areas of the polygons.

To get an accuracy rating for each polygon, the peat polygons were spatially joined with the accuracy rating polygons. In the original dataset, Finland had been divided into five different zones for the accuracy ratings, and accuracies had been calculated separately for each zone. Within each zone two to four accuracy rating values were given, and these values were accompanied with a PDF file including the actual accuracy ratings for the machine learning algorithm used to predict the type of peatland at the given pixel location. Hence, these accuracy rating values were manually extracted from the PDFs and added to the accuracy polygons with the corresponding values. If the accuracy rating was missing for any peat polygons after the merge, they were given an accuracy value of 0.

Besides the accuracy ratings, the peat polygons' areas were used to exclude single peat pixels. As each pixel corresponded to a 10x10m area, the corresponding polygons' areas were 100. Therefore, only polygons with areas of 200 or greater were retained.

These peat polygons could then be joined with the agricultural field parcels from the Finnish Food Authority to obtain peat field parcels. This was done by spatially overlapping the datasets. Separate joins were done with the field parcels for each year from 2020 to 2023. In Figures 4 and 5 are visualizations of this overlap.

Figure 4. Snippet of field parcels

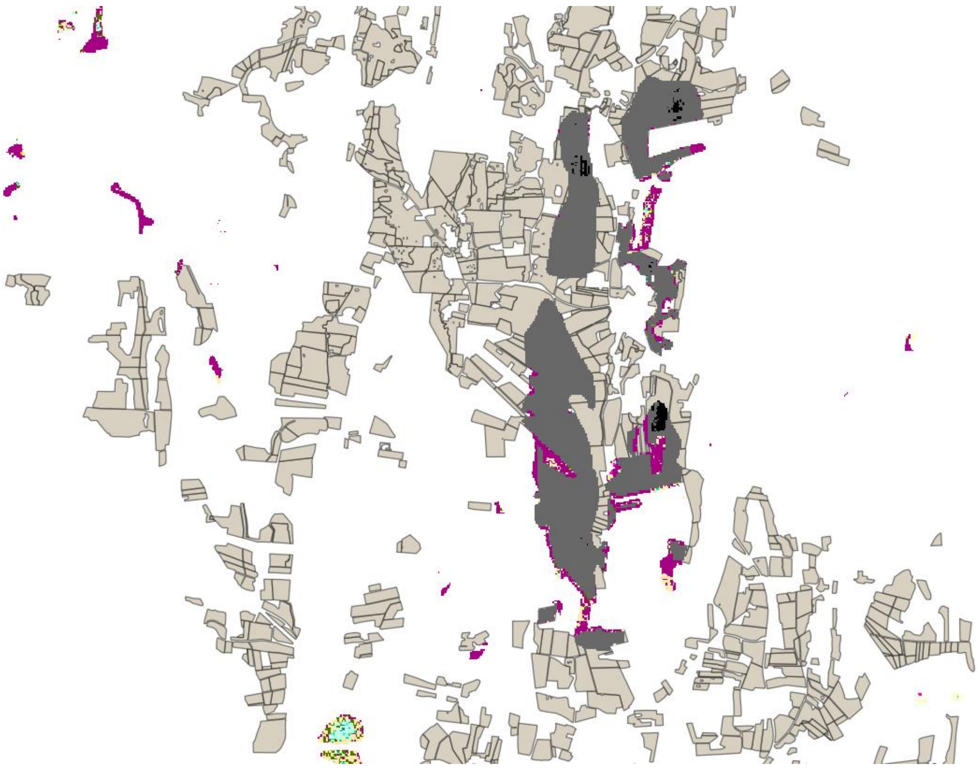


Figure 5. Field parcels’ overlap with drained peatlands

In Figure 5, the dark grey areas correspond to the peatland type of interest: organic soil agricultural fields. As can be seen, the overlapping areas can be of varying proportions of the total field parcel area. Very small peat areas within the field parcels were determined to be insignificant for the parcel. Hence, additional requirements were defined for which parcels to consider peat field parcels: at least 10% of the parcel area needed to be covered by the peat polygon or the overlapping area needed to be at least 1 hectare. The second condition was added as some parcels might be very large in which case a smaller percentage of the total area is still a significantly large area. A hectare was chosen as the limiting value by investigating parcel areas for 2023. 35% of all field parcels had an area of 1 ha or less and therefore peat areas larger than this were considered to be significant.

3.2.2 Peat fields in regions of Finland

Using the resulting information from the previous step, the number of peat field parcels could be calculated. Similarly, the total area of the peat field parcels could be calculated. These statistics were only calculated for the year 2023.

First, the parcel data and the peat parcel data were spatially overlaid with the administrative areas. This allowed for division of parcels between the different regions. Then the peat field polygons and the field parcel polygons were separately dissolved by the parcel identifier number and region. The resulting number of rows in the datasets corresponded to the number of peat fields and field parcels in Finland. Using the resulting geometries' areas, the total areas for both parcel datasets could be calculated by grouping the datasets by region and summing the areas.

Separate dissolved peat parcel datasets were created for each year from 2020 to 2023 by dissolving the peat field parcels by the crop types and the region name. The other columns were aggregated to count the number of appearances of each crop in each region. The resulting datasets were then truncated in a way that only the top five most common crops based on number of occurrences in each region were kept. The yearly datasets were further minimized to include only one row per region by grouping the dataset by the region and aggregating the crops into a single entry with a line break character “\n” between each crop. These four datasets were then merged into one using the region name.

After this, all the information obtained could be merged into a single dataset including each region’s name, the five most common crops for each year between 2020 and 2023, the number of peat fields in the regions in 2023, the number of all field parcels in the regions in 2023, the proportion of peat fields out of all fields in the regions in 2023, and the geometries for each region by merging the desired columns from the dataframes always using the region name column.

3.2.3 Peat fields belonging to real estate

Lastly, the resulting datasets could be overlayed with the cadastral index map to identify fields belonging to real estate as well as to identify real estate with peat fields. This was only done using the most recent data, meaning field parcels from 2023. Only real estate of the type “stead” were included in this analysis to minimize false overlaps with other property types such as roads. This was done by filtering the real estate data frame using the subcategory column.

Next, the real estate data frame was overlapped with the peat field data frame. To prevent incorrect overlaps resulting from spatial variance between the datasets, only fields with 10% or more of their area were within a real estate were included. If the overlapping area was less than 10% of the field’s area, the match was not discarded. However, with this scenario small real estates with only a small part of one field parcel within their property would be excluded. Hence, a second condition was needed. An additional check was added to see if the field area of the overlap was 90% or more of the total field area of the real estate. Overlaps fulfilling either one of the conditions were included in the resulting dataset.

When assigning the real estate to the appropriate regions, spatial overlaps were used again. However, parts of certain real estate resided in different regions. For this reason, all real estate were assigned to the region with the biggest spatial overlap and the duplicate entries in the other regions were excluded.

Next, the total field areas and peat field areas were calculated for each real estate by individually grouping the data frame by the real estate identifier and summing the two types of field areas together separately. Then the peat field proportion was defined by dividing the peat field area by the total field area. After this calculation maximum values of slightly over 1 were achieved. These were still rounded down to 1.0 as small errors can be considered a result of spatial mismatches between the different layers of data.

3.3 Dashboard application design and development

A simple prototype in the form of a Streamlit application could then be developed to display the results of the analysis. This was done iteratively following the principles of design science starting from the step “Define Requirements” and continuing to design and development, demonstration, and evaluation [51].

Before any designing or development started, discussions were had with the case organization to understand their goals and desires for the dashboard. Based on these discussions the main functionalities desired for the solution were identified. These were then turned into an initial list of functional requirements described below in Table 6.

Table 6. Requirements for visualization dashboard

ID	Requirement	Description	Priority
FR-001	Display country-level statistics	The distribution of proportions of peat fields in Finland and the most common crops must be visible on the application.	M
FR-002	Search for a real estate	The application must allow for the lookup of a single or multiple real estate for a closer inspection.	M
FR-003	Display peat fields within a real estate	Map displaying real estate must include peat field parcels within the real estate.	M

Table 6. Requirements for visualization dashboard (continued)

ID	Requirement	Description	Priority
FR-004	Display real estate in region	The application must have a map displaying real estate in the different regions.	M
FR-005	Colour code real estate	The real estate containing peat fields displayed on a map should be colour coded.	S
FR-006	Crop type trends	The application must show evolution of crop types on peat fields over the years available in each region.	M
FR-007	Simple error handling	The application must show an error message if no matches are found when searching for a real estate by id.	M
FR-008	Display all field parcels	When looking for individual real estate, all field parcels belonging to the real estate must be displayed on the map.	M

All these requirements were given the highest priority “must” except for the colour coding of the real estate. It would significantly improve the clarity of the application and provide visual information into the real estate but is not central for the functioning of the application.

Streamlit was the desired development framework as it is rather easy to use, designed for data applications, and often used for data science purposes [55]. A minimal design process was conducted including a plan for the user interface. The application would include three tabs, each focused on a different level of detail to help users identify the level of detail they need. A simplified visual description of the design is shown below in Figure 6.

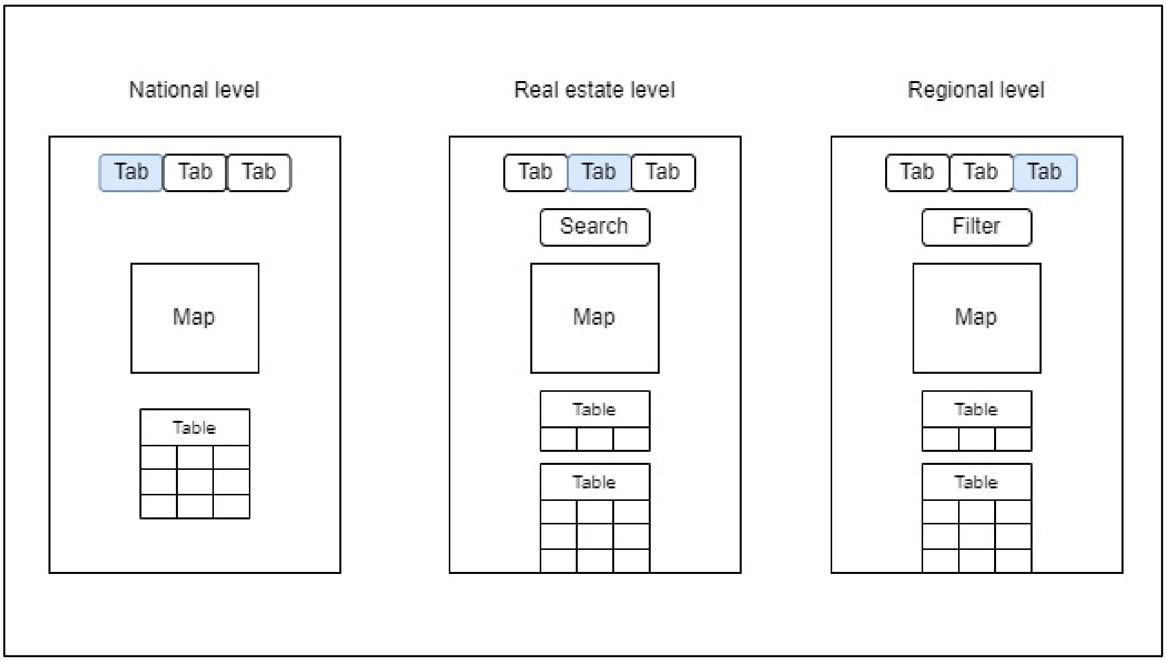


Figure 6. Application design

The first page would have a map that displays the distribution of peat fields in the different regions based on the proportion of peat fields out of all fields. Additional information would be shown below the map, such as the proportion of peat fields and the most common crop in each region.

The second page would display information on a real estate level. Filtering should be added to retrieve information related to a real estate of choice from the real estate, peat field, and field parcel datasets. These three layers of data would then be displayed on a map. Additional information would again be shown below the map, including information about the peat field proportion on the real estate and the crops grown on the field parcels within the real estate.

Finally, the third page would display all real estate with peat fields in the regions. Filtering options would be added to the top of the page to allow for region selection and to limit the real estate shown on the map, for example, based on the peat field proportion. Allowing to limit the data shown was assumed to improve the usability of the application compared to all data being shown.

Based on this design, a first round of development was carried out. Once the first version was developed, it was shown to key stakeholders at the case organization. In this work, this step corresponded to the demonstration stage of design science where the artefact is used in an illustrative case to test its feasibility as well as the evaluation stage where its ability to fulfil the defined requirements are evaluated [51]. After the demonstration feedback was received. The previous requirements were satisfactorily fulfilled but some additional requirements were identified based on this session.

Table 7. New requirements

ID	Requirement	Description	Priority
FR-009	Divide regional data to groups	On the region page, display pre-defined options “at least 0.5”, “less than 0.5”, and “all” for each region.	S
FR-010	Pre-selected peat field proportion	On the region page, “at least 0.5” should be automatically selected.	S
FR-011	Display crops for all years	Crops for all years 2020-2023 should be displayed on the individual real estate page.	S

All the new requirements shown in Table 7 were given the priority “should” as they significantly improve the application, but they are not central to its functioning. The new requirement to display crops for each year in real estate was added to help track perennial cultivation and changes. The other two requirements would help users view real estate on the regional level, especially improving the usability and responsiveness of the application.

Additional changes to be made were identified for the user interface. As some crops names had been simplified from the years 2020-2022 for 2023, simplifications were made for the 10 most common crop names where applicable. The crops on the page displaying information about all of Finland were switched to show the most recent ones from 2023. The tabs were also re-arranged to go from the highest level to the lowest level, in other words from national level to regional level to real estate level.

3.4 SusAF

After development, the evaluation process of the solution was continued with an analysis of its impact on the environment it operates in. This was done following the SuSAF framework. No key stakeholders beside the researcher were involved at this step. The capture and analysis phase were followed according to instructions in the workbook [54] step by step.

The process started with brainstorming. All sustainability dimensions were covered individually through the questions available in the workbook. Based on these questions, over 60 possible effects were identified. Out of these, 23 effects were determined to be the most concrete and important based on the expected impact and likelihood. They were then placed on a prioritization matrix according to the likelihood of the effect realising as well as the impact of the effect should it realise. This helped with identifying the most important effects as they were placed in the upper right corner where the likelihood and impact were high.

The most important effects were then placed on the SusAD. The diagram allows for visualization of the distribution of the effects in the different dimensions of sustainability, the connections between the effects, and their order of impact as defined in [54]: immediate, enabling and structural.

In the immediate order of effects, were placed effects directly related to the production, operation and usage of the system. Effects resulting from the usage and potentially changed behaviours were included in the enabling order. Lastly, the third-order, structural, included effects manifesting in, for example, politics and social norms through structural changes by the long-term usage of the solution.

After the most important effects were added to the correct sections of the diagram, effects resulting from them or effects leading to them were determined from the remaining effects. Some effects from the matrix were included in multiple dimensions or broken down into two effects of different order.

A full synthesis was not developed within this research, but some opportunities and risks and corresponding actions were discovered. They will be presented in Chapter 4 and further discussed in Chapter 5.

4 Results

The following sections will describe the results of the research. First, some statistics related to peat fields and real estate in Finland will be presented. Then, a description of the developed prototype will be covered. Lastly, the results of the SusAF analysis will be shown.

4.1 Peat fields in Finland

During the data processing, a peat field parcel was defined as a parcel where the peat coverage was at least 10% of the total parcel area or the peat area was at least 1 hectare. With this definition 151 474 peat field parcels were identified in Finland for the year 2023. Out of these only 6.99% were completely covered by peat, meaning the peat area coverage was at least 99% of the parcel area. In total, the peat coverage of field parcels in Finland was calculated to be 14.17%. However, there was significant variance between the different regions of Finland.

As can be seen from Table 8 on the next page, differences arise when comparing the proportion of peat fields based on the number of parcels and the field area. In some cases, the proportion of peat fields is significantly different, for example, in Lapland the difference is over 7 percentage points. Nevertheless, at the national level the differences are balanced out.

Table 8. Peatlands in the different regions of Finland

Region	Peatlands based on number of parcels (%)	Peatlands based on area (%)
Åland	2.70	2.00
South Karelia	16.33	15.46
South Ostrobothnia	21.88	19.23
South Savo	11.87	11.69
Kainuu	32.85	38.59
Kanta-Häme	12.13	8.92
Central Ostrobothnia	30.91	32.79
Central Finland	13.56	13.67
Kymenlaakso	8.26	6.03
Lapland	34.03	41.73
Pirkanmaa	9.66	8.47
Ostrobothnia	9.66	9.62
North Karelia	18.46	19.13
North Ostrobothnia	29.76	31.27
North Savo	16.96	16.35
Päijät-Häme	7.26	5.38
Satakunta	14.06	11.31
Uusimaa	4.94	3.49
Southwest Finland	4.26	3.00
Finland total	**15.49**	**14.17**

Out of the regions Central Ostrobothnia, Kainuu, and Lapland had the most peat fields with North Ostrobothnia not far behind. On the other end, Åland, Uusimaa, and Southwest Finland had the least peat fields. The distribution is displayed on a map below in Figure 7.

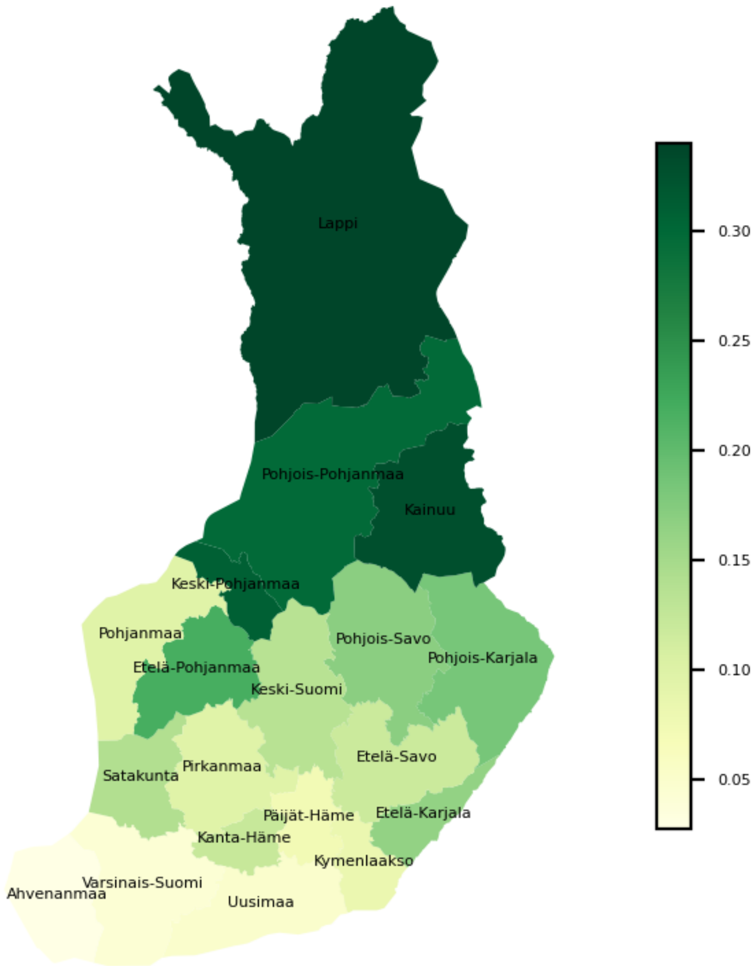


Figure 7. Distribution of peat fields in Finland based on area

On this map the proportion displayed is the proportion based on number of field parcels. Overall, there is a trend of a bigger percentage of peat fields in the north than in the south.

Regarding the proportion of peat fields belonging to real estate, a comparable trend was established. The distribution between the different regions was similar to the overall peat field proportions. The mean proportions of peat fields belonging to real estate in each region is shown below in Table 9 in increasing order.

Table 9. Mean proportion of peat fields in real estate per region

Region	Mean proportion
Ahvenanmaa	0.23
Southwest Finland	0.27
Uusimaa	0.29
Päijät-Häme	0.30
Kymenlaakso	0.33
Pirkanmaa	0.36
North Savo	0.37
South Savo	0.37
Kanta-Häme	0.37
Central Finland	0.38
Ostrobothnia	0.40
South Karelia	0.42
Satakunta	0.43
North Karelia	0.44
South Ostrobothnia	0.46
North Ostrobothnia	0.50
Central Ostrobothnia	0.51
Kainuu	0.54
Lapland	0.55

Noticeably Ahvenanmaa, Southwest Finland, and Uusimaa are again at the lower end whereas Central Ostrobothnia, Kainuu, and Lapland have the highest mean proportions. In the top four regions the mean peat proportion is at least 50% whereas the lowest mean values are around 25%. The maximum peat proportion for all regions was 100% and the minimum proportion close to 0, amounts in the order of 10^{-6} to 10^{-11}.

The proportion of peat fields in real estate could also be analysed on the national level. The distribution is displayed below in Figure 8. On the graph, the data is divided into 40 bins, each bin shown as a single pillar. The height of the pillar corresponds to the number of real estate with the given range of peat field proportions.

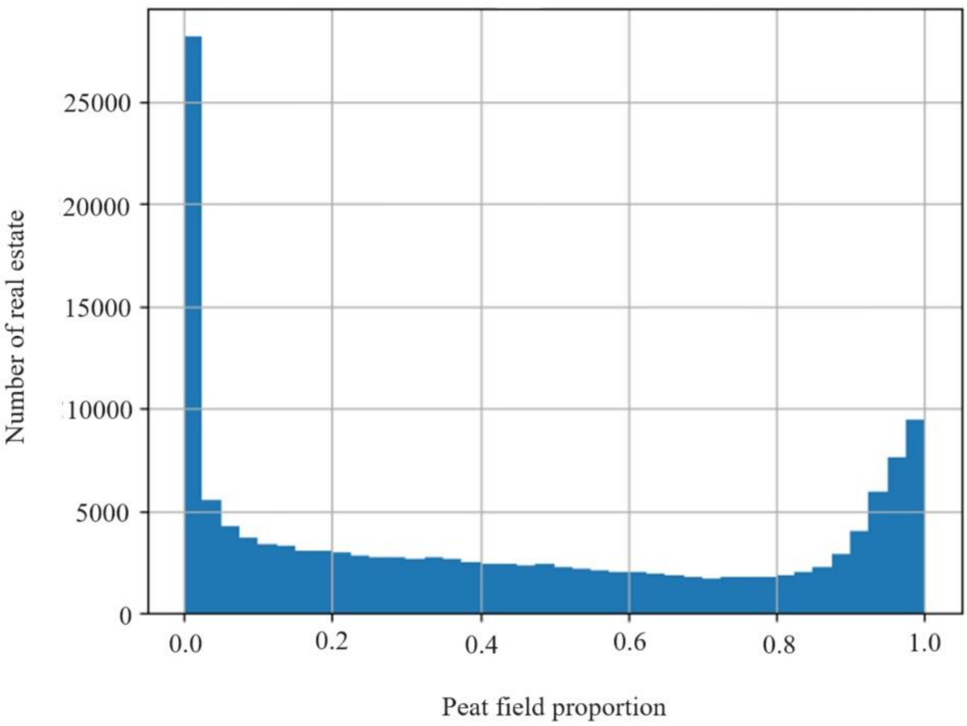


Figure 8. Proportion of peat fields in real estate

The graph highlights the most common proportions of peat fields on the very left. Each pillar of the graph corresponds to a proportion of 0.025. Hence, the first pillar represents proportions between 0 and 0.025. 19% of all Finnish real estate containing peat fields belong in this category. From there on the group sizes decrease until about 0.7 when the number of real estate starts increasing again. The second biggest group is on the other end of the graph from 0.975 to 1 with 6.5% of real estate in this category.

4.2 Crop types

Although the peat fields contained multiple different types of crops, some were noticeably more common than others. A representation of the most and least common crop types nationwide per year was not added to the dashboard application, but it was investigated during the data analysis. The full table can be found from Appendix 1.

The most common crop "perennial grass (dry hay, silage, soilage)" was the same for all years until 2023 when it became "fodder grass" instead due to a naming change. Other common crops from year to year included oat, fodder barley, and nature management grass and fields. The least common crops were more varied due to the low numbers of occurrence and a change by one appearance removing or adding a crop to the list.

Similar to the nationwide statistics, the most common crop types were calculated for each region. The results for 2023 are presented below. If multiple regions had the same top crops in the same order, they were grouped together. The least common crop types were not checked in this analysis.

Table 10. The top 5 most common crops per region in 2023

Region	Crops
Åland	Fodder grass Oat Potatoes for food industry Green fallow (grass and meadow) Fodder barley
South Karelia, Kanta-Häme	Fodder grass Nature management grass Oat Fodder barley Spring wheat
South Ostrobothnia	Fodder grass Oat Fodder barley Nature management grass Polymorphism crops, game
South Savo	Fodder grass Oat Nature management grass Fodder barley Grazing grass
Ostrobothnia, North Ostrobothnia, Central Ostrobothnia	Fodder grass Fodder barley Oat Nature management grass Polymorphism crops, game
Kymenlaakso	Fodder grass Oat Landscape grass Spring wheat Fodder barley

Table 10. The top 5 most common crops per region in 2023 (continued)

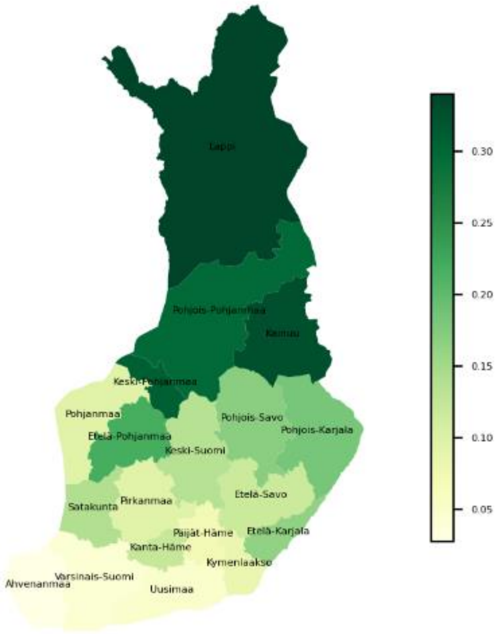
Region	Crops
Pirkanmaa	Fodder grass Oat Nature management grass Fodder barley Spring wheat
North Savo	Fodder grass Fodder barley Nature management grass Oat Green fallow (grass and meadow)
Päijät-Häme	Fodder grass Nature management grass Oat Fodder barley Malting barley
Satakunta	Fodder grass Oat Fodder barley Nature management grass Spring wheat
Uusimaa	Fodder grass Nature management grass Oat Spring wheat Green fallow (grass and meadow)
Southwest Finland	Oat Fodder barley Nature management grass Fodder grass Spring wheat
Kainuu	Fodder grass Green fallow (grass and meadow) Nature management grass Fodder barley Grazing grass
Central Finland	Fodder grass Nature management grass Fodder barley Oat Polymorphism crops, game
North Karelia	Fodder grass Oat Fodder barley Green fallow (grass and meadow) Nature management grass
Lapland	Fodder grass Nature management grass Green fallow (grass and meadow) Polymorphism crops, game Grazing grass

When looking at the top five crop types, mostly the same ones are repeated in all regions with grass being the most common type in all regions except for Southwest Finland where oat was the most common type. However, even there grass was among the top 5 crops and oat was among the top 5 crops in all other regions except for Lapland. Different types of grass and grains took almost all top 5 spots within Finland.

4.3 Dashboard application

After the data analysis was completed, the visualization dashboard was developed to display the results. Based on the requirements described in Chapter 3, three tabs with varying functionality were implemented fulfilling all the requirements. A full version of this application would run in a cloud environment, but the prototype was developed locally.

The first tab provided an overview of peat fields in Finland. A snippet of this page is displayed below in Figure 9.



Maakunta	Turvepelto-osuus	Yleisin kasvi
Ahvenanmaa	0.0270	Rehunurmi
Etelä-Karjala	0.1633	Rehunurmi
Etelä-Pohjanmaa	0.2188	Rehunurmi

Figure 9. First page displaying country-level information

This page includes a static map with the regions of Finland colour-coded based on the proportion of peat fields out of all field parcels in the region. Additional information about the regions is displayed in a table below the map: the exact proportion value and the most common crop in the peat fields in 2023 for each region.

A second tab displayed all real estate containing peat fields in a given region. A dropdown selection was added to the top of the page and below it a peat proportion selection and a map displaying the matching selection, as displayed below in Figure 10.

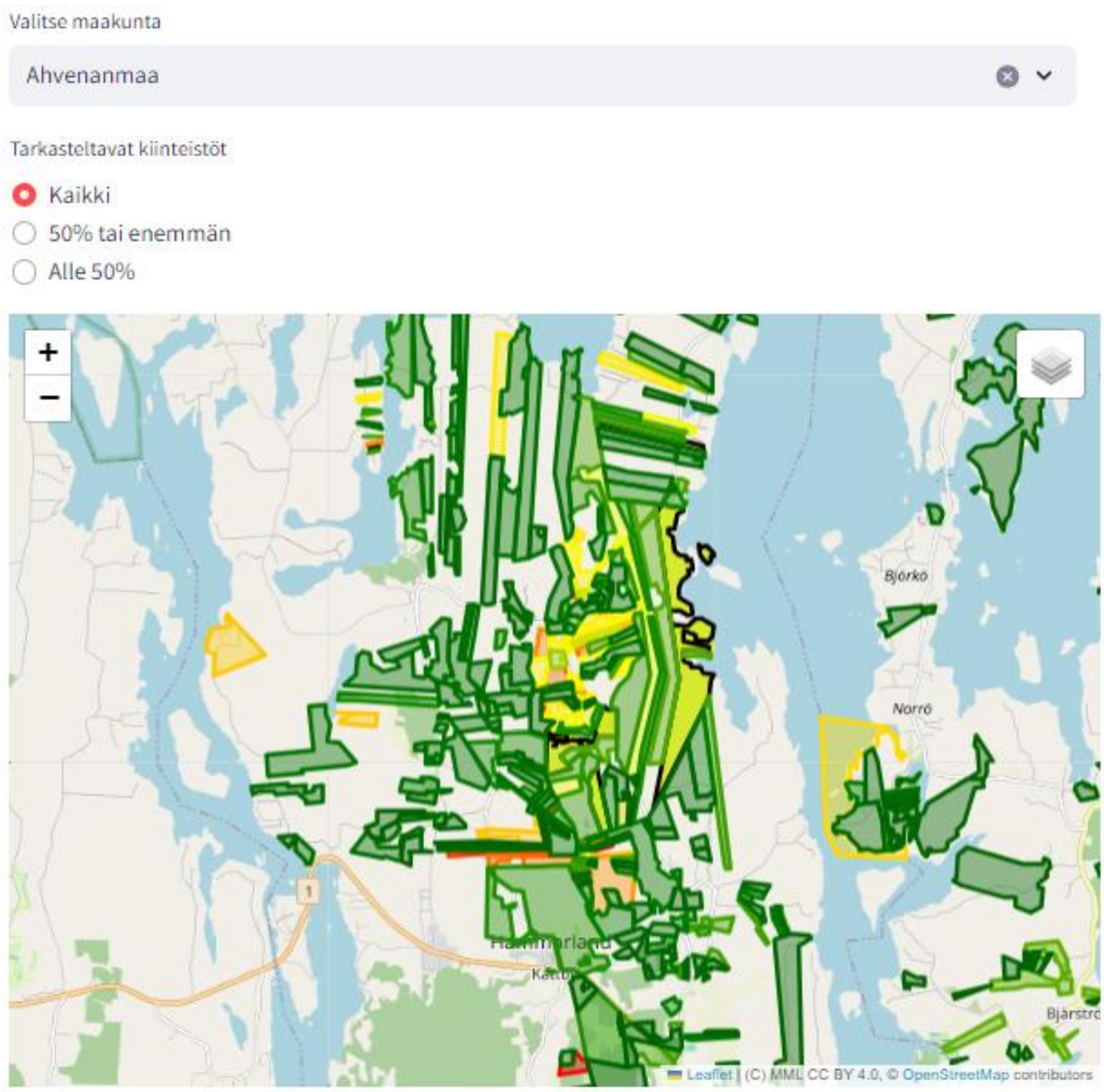


Figure 10. Snippet of all real estate in Åland

When hovering over a real estate, it becomes highlighted with the same colour and its border turns black, as is the case for the real estate displayed as light green towards the centre of the map in Figure 13. In this implementation, the colour is determined by the proportion of peat fields: the greener the colour the smaller the proportion, and the redder the colour the greater the proportion.

Below the map was displayed additional information about the region in two tables. The first table contained the number of peat field parcels and all field parcels in the region as well as the proportion of peat fields parcels both based on count and area. The second table showed the top 5 crops in the region between 2020 and 2023.

Peltojen lukumäärä

Turvepeltojen lukumäärä	Peltolohkojen lukumäärä	Turvepelto-osuus lukumäärästä	Turvepelto-osuus pinta-alasta
334	12365	0.0270	0.0200

Yleisimmät kasvit

Yleisimmät kasvit 2020	Yleisimmät kasvit 2021	Yleisimmät kasvit 2022	Yleisimmät kasvit 2023
Rehunurmi Ruokateollisuusperuna Viherkesanto (nurmi ja niitty) Kaura Rehuohra	Rehunurmi Viherkesanto (nurmi ja niitty) Ruokateollisuusperuna Rehuohra Kaura	Rehunurmi Viherkesanto (nurmi ja niitty) Ruokateollisuusperuna Rehuohra Kaura	Rehunurmi Kaura Ruokateollisuusperuna Viherkesanto (nurmi ja niitty) Rehuohra

Figure 11. Additional information for Åland

As shown in Figure 11, minimal styling was added to the tables. This was in an attempt to keep them as clear as possible.

The third page displayed information about particular real estate. Users can input the real estate identifier into an input field at the top of the page. If the identifier is not found from the dataset, an error message is displayed stating that the identifier was not found, shown below in Figure 12.

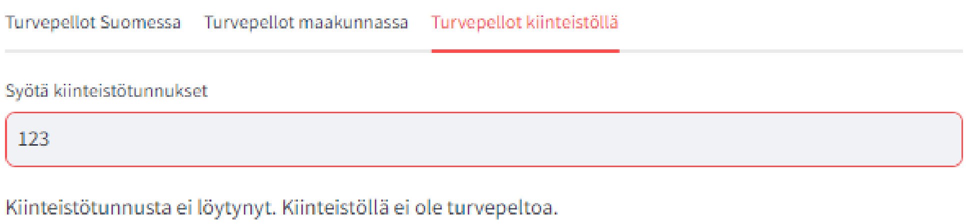


Figure 12. Error message for missing real estate identifiers

As can be seen from the input in Figure 12, the format of the input was not restricted at this point. In the message it is assumed, that identifiers are not found because real estate with the identifier does not contain any peat fields. If the user inputs at least one identifier that is found from the dataset, the results are displayed, as shown below in Figure 13.

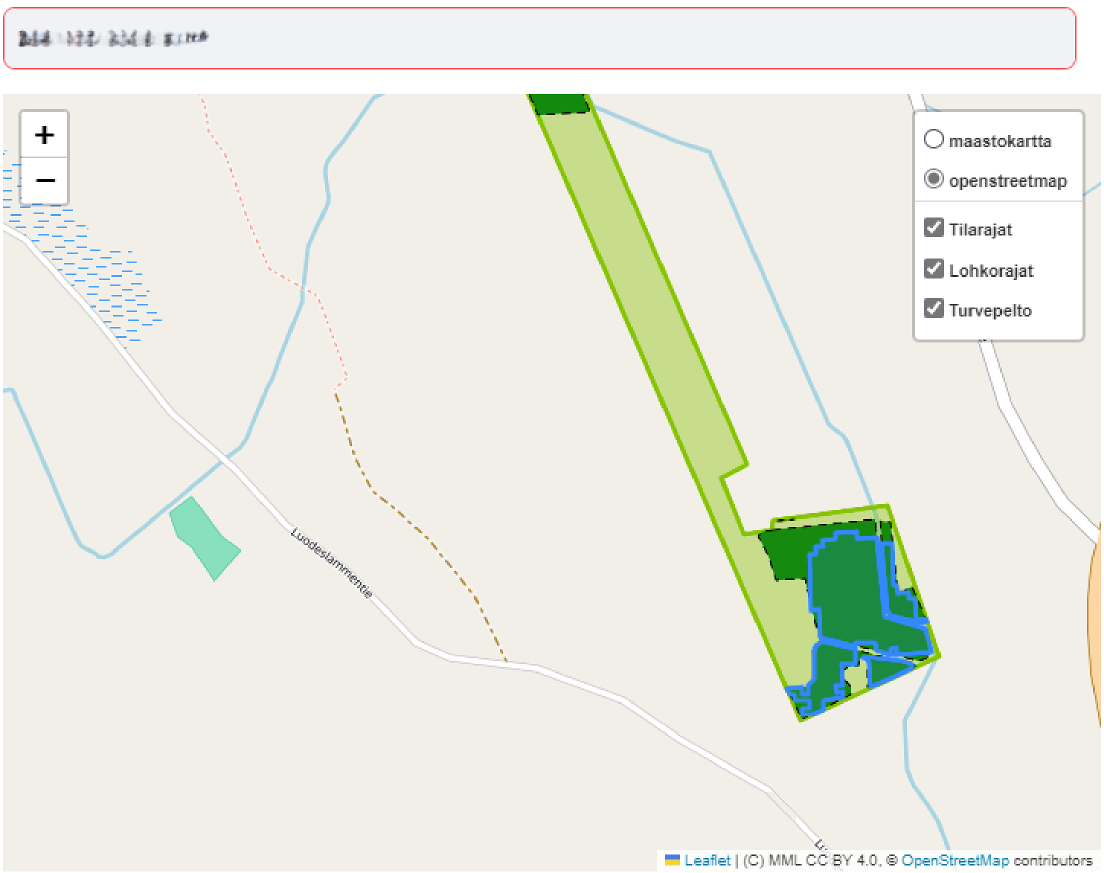


Figure 13. Viewing information about a real estate on a map

Three layers can be chosen to be displayed: the real estate area, the field parcels within the real estate, and the peat field areas within the real estate. The real estate area is drawn in a light green colour whereas the field parcels within the real estate are drawn with a darker green colour and a black dashed outline. The peat area within the parcels is drawn in blue. One or more real estate can be searched for and displayed at once to allow for viewing of entities consisting of multiple separate uniquely identified real estate. Below the map, more information, shown in Figure 14, about the real estate was displayed.

Kiinteistön tiedot

Kiinteistötunnus	Turvepeltoala	Peltoala	Turpeen osuus peltoalasta
[illegible]	20.8229	26.2538	0.7931

Kasvulohkotiedot kiinteistöittäin

Kiinteistötunnus	Peruslohkotunnus	Kasvilaji 2020	Kasvilaji 2021	Kasvilaji 2022	Kasvilaji 2023
[illegible]	[illegible]	Suojavyöhyke (sitoumus alkaen 2015)	Suojavyöhyke (sitoumus alkaen 2015)	Suojavyöhyke (sitoumus alkaen 2015)	Rehuohra
[illegible]	[illegible]	Rehunurmi	Rehunurmi	Rehunurmi	Rehunurmi
[illegible]	[illegible]	Rehuohra	Rehuohra	Rehunurmi	Rehunurmi
[illegible]	[illegible]	Suojavyöhyke (sitoumus alkaen 2015)	Suojavyöhyke (sitoumus alkaen 2015)	Suojavyöhyke (sitoumus alkaen 2015)	Rehuohra
[illegible]	[illegible]	Rehunurmi	Rehunurmi	Rehunurmi	Rehunurmi
[illegible]	[illegible]	Rehunurmi	Rehunurmi	Rehunurmi	Rehunurmi
[illegible]	[illegible]	Rehunurmi	Rehunurmi	Rehuohra	Vihantavilja (viljaseos)
[illegible]	[illegible]	Rehuohra	Rehuohra	Rehunurmi	Rehunurmi

Figure 14. Additional information about the chosen real estate

For each real estate, both the field area as well as the peat field area were displayed in a table. Additionally, the proportion of the peat field area out of the total field area was shown. Below this information, a separate table was added to list the field parcels and the corresponding crop each year. In the case of multiple real estate, all real estate entities and field parcels would be shown in a single table.

4.4 Sustainability implications

The solution's sustainability impacts were evaluated using the SusAF framework as described in Section 3.4. The analysis results are presented below starting with the prioritization matrix (Figure 15). Positive effects were written on blue cards and negative effects on orange cards. In this scenario, it was assumed that the solution would be available for anyone wishing to access it.

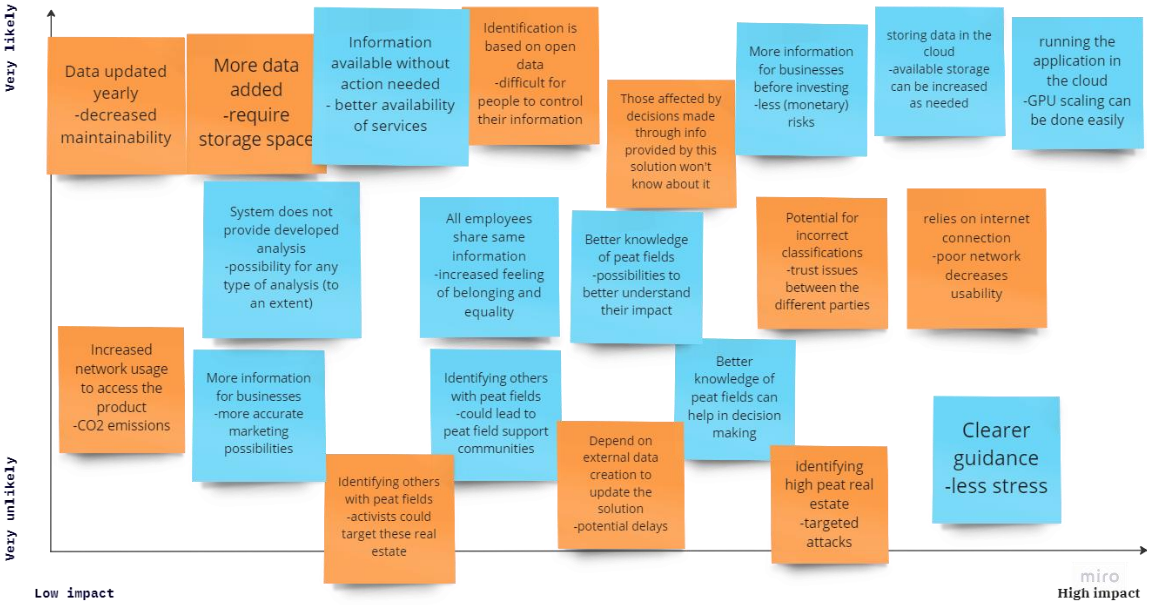


Figure 15. Prioritization matrix for identified effects

The most important effects are in the upper right corner with high likelihood and high impact. The positive effects there are related to the usability of the application through scalability and the possibilities related to more important information being available. On the negative side, the effects are more varying: network connections can decrease the usability of the application, problems with the peat field classifications could lead to trust issues, and the people affected by the decisions made through the usage of this application will not know about the process and how it has affected them. The SusAD diagram for visualizing the connections between the most important and related effects is shown below in Figure 16.

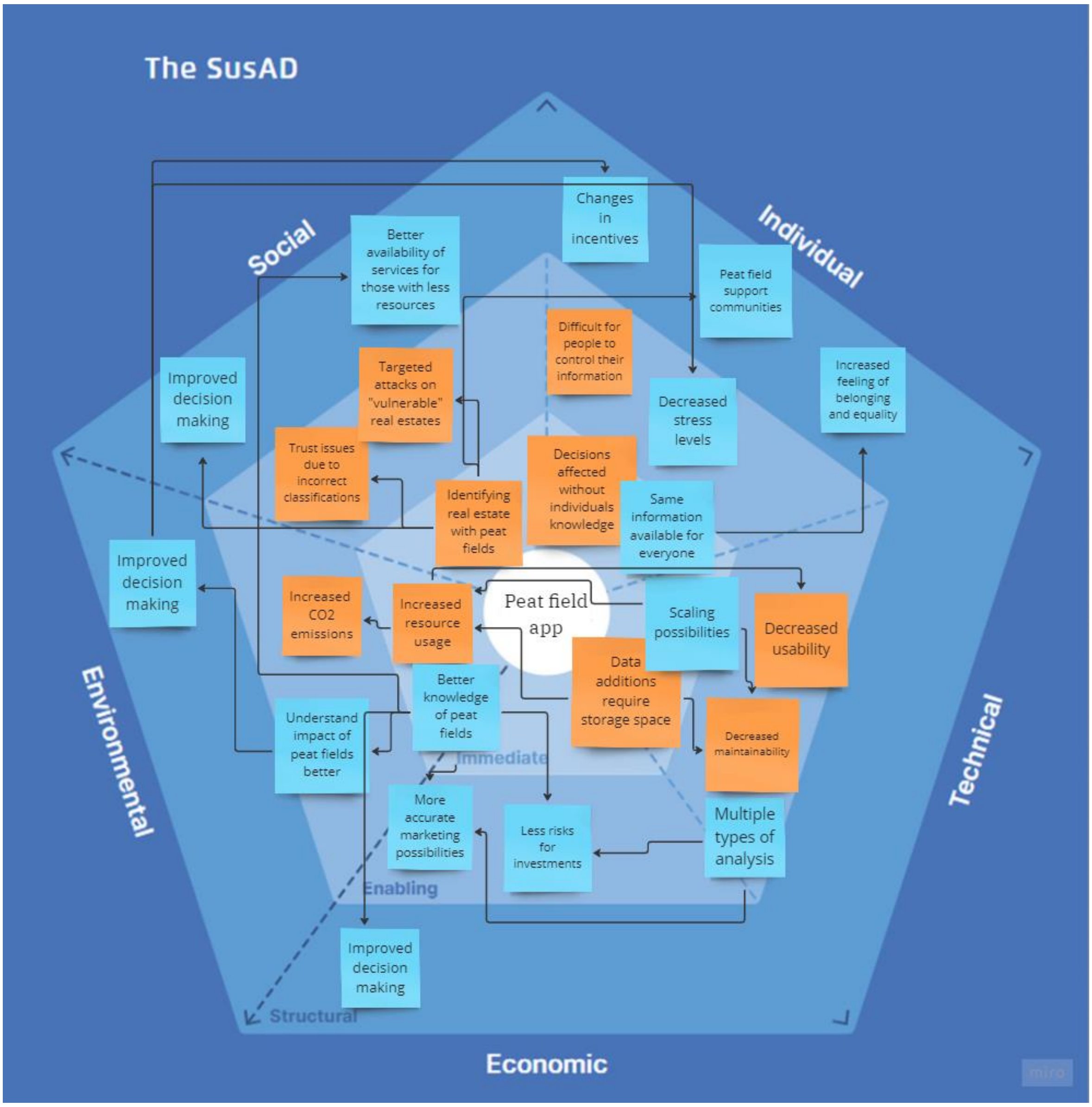


Figure 16. SuSad diagram for sustainability effects. Template source: [54]

A large part of the identified effects are related to more information being available, which can have both positive and negative impacts. One of the negative impacts is the inevitable increase in resource utilization as a new solution and new data are being introduced. However, as the application is a single-page web application the additional resource requirements can be considered low. Additionally, most of the increases are related to energy and electricity usage, and ensuring the electricity used is produced in a climate neutral manner would not lead to an increase in emissions.

On the positive side, the increased availability of information could lead to a better understanding of peat fields and improved decision making. These decisions could then have positive impacts on different aspects of the environment, such as the climate and biodiversity. The decisions could also help guide incentives for peat fields which could be a positive or a negative impact but, in this case, it is assumed that it would be done in the best interests of the individuals.

When it comes to the technical dimension of sustainability, the most important factors for this solution are related to scalability. As new data is created, the need for efficient storing and displaying of this data will become more evident. Deploying the application in a cloud environment could provide efficient scaling options and would have a positive impact on the usability of the application as the amount of data and the number of users increase.

From a maintainability point of view, it could make sense to have more of the data processing and combining happen in the application or in a backend server at runtime. Having only more processed data available limits the possibilities of what can be done in the application and can therefore hinder its maintainability.

On the individual side, both positive and negative impacts can arise. On the one hand, the tool could help spread the same amount of information between all user which could increase their feelings of belonging. Although not the intended usage of the solution, it could also help users identify other peat field owners which could result in like-minded groups being formed. On the other hand, as the solution is based on open data, it is difficult for individuals to control their information and the decisions made by others using the application.

Although the open data makes it difficult to control one’s information, it also means that individuals do not need to take any action for their information to be available. This could be beneficial for those that do not have the resources or knowledge to provide such information as potentially more services could be provided for them without their input. This information could also be useful for companies that might want to target their marketing to certain types of real estate. Furthermore, this information could be of interest to those planning on investing in certain real estate as it could lower their risks in cases where peat fields would have an impact on their investment.

The provided solution would allow for multiple types of analysis to be conducted using the information available. Although this allows for many positive effects, it could also be used without good intentions. Albeit unlikely, extreme climate activists could identify real estate with peat fields using the solution and target them in an attempt to make a change. Should it happen, it would cause stress on the society and the individuals.

On the other hand, the solution could also alleviate individuals stress levels. If more information about peat fields is obtained and used to positively guide decision making and regulations, it can result in clearer guidelines and action points for these individuals. This could help them understand what is needed and expected of them.

5 Discussion

Overall, combining and visualizing information from the different geospatial data sources was successful. Like previous studies, a number of different datasets were combined through the use of GIS and spatial joins. Although the datasets were from different sources and different providers, they worked well together with minor spatial differences between them.

Conducting the research as a case study focusing on Finland as the case area also provided similar results to some previous studies: in [21] a similar trend of proportions of peat fields in the different regions of Finland was discovered. The calculated total peat coverage of agricultural fields in Finland was also in an acceptable range, as it has previously been estimated to be around 11% [19] whereas this study resulted in a value of approximately 14%. The geographical limit was useful in limiting the amount of data as well as the types of data needed without decreasing the meaningfulness of the work. However, as the methods and results have been analysed only for the case area, the applicability should be investigated with other cases.

The current process for peat field identification is dependent on the accuracy of the available datasets. When it comes to the peatland data, the accuracy ratings for the organic soil agricultural fields were mostly above the overall average at around 80-90% but also 0% at times. However, these 0% areas were areas where only 1-3 data points were available for identification, meaning a lower overall percentage of peat fields. Hence, the reliability of the used data can be considered high. Regardless, a few known non-organic field parcels were discovered to have been identified as peat fields. This was not dependent on the data processing but instead a false identification in the original data. These types of false positives are impossible to omit as it would require ground truth values for each parcel. This would also mean that the identification process using geospatial data would not have been needed in the first place.

The accuracy values of organic soil agricultural fields were missing from two accuracy PDFs although multiple pixel points had the accuracy value assigned to them. A threshold value was also used when deciding which points to include in the peat field analysis which means that only the most reliable values were used but this does not completely eliminate the risk of false identifications. Similarly, false negatives could result from the less reliable values being left out.

Additional unreliability could result from the peat field definition used when identifying them. Only field parcels with a great enough peat coverage were considered peat fields and only groups of peat pixels were taken into consideration during the identification process. Depending on the desired usage of the data analysis, this could potentially skew the results.

The combination of data from different sources creates marginal errors as well. The peatland data seems to be based on data from 2021 meaning that when combining it with data related to real estate or field parcels from other years, the landscape might have changed, or field parcels might have been added to or removed from the registry between the years. In the Finnish conditions new peat is formed at a rate of about 1mm per year [21] and the peat layers get thinner on average at a rate of 1.2cm per year [21]. Therefore, the changes within such small amount of time can be considered negligible. Similarly, minor adjustments needed to be made when spatially overlaying the datasets as they mostly do not overlap perfectly. This was especially evident when connecting field parcels to real estate.

The identification of field parcels belonging to real estate could be a prospect for further investigation. Upon closer inspection it was determined that when filtering through the real estate to avoid including roads and other incorrect types of real estate when farms are the interesting type, some roads were categorized as “stead” in the dataset. As these entities were classified as “properties” and with the subcategory “stead”, it was not possible to separate them from the farm real estate. Even with the additional requirement of certain field area within the real estate some roads remained in the dataset. Most uninteresting real estates were still filtered out with the chosen parameters and considering the amount of data used, the few mismatches do not significantly contribute to the results.

Improvements could also be investigated for the developed prototype. The possibility to display all field parcels belonging to the individual real estate on the regional map was investigated during the development phase but due to the high number of fields and real estate in certain regions, the idea was disregarded during this round of development. The size of the datasets would have required more memory usage and noticeably slowed down the service, hence, the cons were considered to outweigh the pros of having the visualization but if more powerful visualization tools would be investigated, this is a potential visualization to come back to. The balance between responsiveness and maintainability should be carefully evaluated and a best balance should be identified. Some functionalities and packages were unavailable for use at this time due to company policies and therefore could not be properly investigated within this research.

As a result of the data analysis and software development processes, a working prototype displaying information about peat fields in Finland could be delivered to the case organization. This solution would allow them to make interpretations of the statistics according to their desires. Discussions and evaluations of the prototype were held with some key stakeholders of the case organization, but no user testing was conducted with potential users of the solution to get exact usability feedback. This could be a future topic of investigation.

The results of the data analysis can be used to identify individual real estate containing peat fields. This information could be useful in directing regulations and improving sustainability practices. Site-specific policies are needed to target multiple goals related to for example, biodiversity and soil productivity, to diminish negative impacts on the climate [33] and utilizing this type of data could be useful to identify the targets for such policies. However, attention should be paid to how to implement activities such as the rewetting of peatlands in a socially just way without causing unnecessary stress on the individuals owning the lands.

Similarly, the results could be used to investigate the importance of peat fields during dry seasons. As the soil is less dry it can be hypothesized that these fields could provide a better or more secure yield during these times. With the help of the identification conducted in this work, this could be further investigated.

As noted earlier, it was discovered that the most common crops grown on the peat field parcels were different types of grass. As mentioned in [56], grass and especially multiyear grass cover is a good option for peat fields in preventing GHG emissions. Hence, this information is a positive discovery for slowing down climate change.

On the other hand, grass is not used as food by humans, so these fields do not appear to significantly contribute to food availability and security. Nevertheless, oat was still one of the most common crop types in almost all regions, and other human consumable crops could be found within the top 10 crop types nationwide. This could be a topic for further research as the information currently available is not sufficient for drawing significant conclusions.

For example, having the yield levels of these fields available would be beneficial for achieving more insights into the importance of these fields.

Similarly, at this point attention was not paid to whether the peat fields were drained or undrained. As the drainage of peatlands is problematic for GHG emissions and the restoration demands are aimed at the drained fields, the distribution and impacts of these types of peat fields is another topic that could be investigated in the future and could change some impacts of the solution.

Another direction of further research could be a greater focus on the real estate. As identification of peat field within real estate has not been possible in the past, the results of this research provide possible baseline information for such works. For example, the proportion of real estate containing peat fields out of all real estate was not checked within this work, but it could be an interesting topic of further research to investigate the differences in the different areas. Similarly, as the number of real estate got greater when the proportion of peat fields approached 1, investigating this trend and the reasons behind it could be a potential research focus in the future.

6 Conclusions

Understanding the state of peat fields is important for climate change mitigation actions. These actions are needed to ensure the wellbeing of the environment and to reach the EU's goals. Besides authorities behind regulations, individuals and companies can also benefit from improved peat field knowledge. Thus, this thesis investigated the possibilities of geospatial data in providing more information.

To identify actions required for the investigation, collaboration with a case organization took place. Their current sustainability analysis practices were explored through interviews and needs for improvement were identified. Based on the interviews and conversations with key stakeholders from the company, the directions and goals for this research were identified.

Further directions were identified through a structured literature review that provided insights into the current state of geospatial data usage in peatland related analysis. A good understanding of the global distribution of peatlands has already been achieved but little attention has been paid to the distribution and qualities of peat fields used for agricultural purposes.

The results of this research were achieved by analysing four types of geospatial data covering Finland: peatlands, real estate, regions, and agricultural field parcels. For this analysis a peat field parcel was defined as a parcel that had at least a 10% peat coverage or that had a peat coverage equal to or greater than 1 hectare. Using this definition, a 14.17% peat coverage of field parcel areas in Finland was calculated. The proportion of peat fields was highest in northern Finland and lowest in southern Finland, corresponding to previous studies on Finnish peatlands.

In addition to the distribution of peat fields, the varieties of cultivated crops were identified. Some common crops included human grade varieties, such as oat, but others were directed towards livestock, such as grazing areas. This information could be used for further analysis on the importance of these fields in food production.

A prototype of a dashboard application was implemented for displaying the results of the data analysis. The solution was approved by the key stakeholders of the case organization suggesting the potential of the data in use. However, user testing was not conducted within this research. Testing should be done and the solution's potential for a variety of sustainability analysis should be further investigated to fully understand its potential in providing meaningful information as well as its impact on the different dimensions of sustainability.

Although the intention of the solution is to provide more information and to have a positive impact on sustainability, potential negative impacts were also identified through the SusAF analysis. On the one hand, the information has the potential to help with sustainability initiatives, such as the EU nature restoration law requiring drained peatlands to be rewetted, and can be useful for companies when making financial assessments. On the other hand, these possibilities can cause stress on the individuals that own peat fields if they are not implemented in a fair manner.

Overall, the insights achieved with this research correlate with previous findings and expand on them. However, this was the first time that peat fields within real estate had been investigated. The results were identified to have potential for both positive and negative long-term impacts on the different dimensions of sustainability. Further studies on their usability and impact on a bigger scale should be investigated to better understand the potential implications.

References

[1] UNEP, 2022. Global Peatlands Assessment – The State of the World's Peatlands: Evidence for action toward the conservation, restoration, and sustainable management of peatlands. Main Report. United Nations Environment Programme, Nairobi: Global Peatlands Initiative. [online]. [Cited 31.7.2024]. Available: https://www.unep.org/resources/global-peatlands-assessment-2022.

[2] Rahman, A., Afroz, R. & Yahya, A., 2012. *The peat swamp productivity, trafficability and mechanization*, 1st ed. New York: Nova Science Publishers.

[3] Carlson, K. M., Gerber, J. S., Mueller, N. D., Herrero, M., MacDonald, G. K., Brauman, K. A., Havlik, P., O'Connell, C. S., Johnson, J. A. & Saatchi, S., 2017. Greenhouse gas emissions intensity of global croplands. *Nature Climate Change*, vol 7. [online]. [Cited 13.4.2024]. Available at: https://doi.org/10.1038/nclimate3158

[4] Gerin, S., Vekuri, H., Liimatainen, M., Tuovinen, J.-P., Kekkonen, J., Kulmala, L., Laurila, T., Linkosalmi, M., Liski, J., Joki-Tokola, E. & Lohila, A., 2023. Two contrasting years of continuous N2O and CO2 fluxes on a shallow-peated drained agricultural boreal peatland. *Agricultural and Forest Meteorology*, vol 341. [online]. [Cited 14.4.2024]. Available: https://doi.org/10.1016/j.agrformet.2023.109630.

[5] Tian, T., Zhang, Y. & Mei, Y., 2022. Intelligent analysis of precision marketing of green agricultural products based on big data and GIS. *Earth science informatics*, vol 15(3). [online]. [Cites 8.4.2024]. Available: https://doi.org/10.1007/s12145-022-00789-5.

[6] IBM, 2021. What is Geospatial Data? [website]. [Cited 7.6.2024]. Available: https://www.ibm.com/topics/geospatial-data.

[7] Xu, J., Morris, P. J., Liu, J. & Holden, J., 2018. PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis. *Catena*, vol 160. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1016/j.catena.2017.09.010.

[8] Saha, S. Das, J. & Halder, S. 2023. *Nexus Between GIScience and Sustainable Agriculture*. In Das, J., Halder, S. (eds) Advancement of GI-Science and Sustainable Agriculture. GIScience and Geo-environmental Modelling. Springer, Cham. [E-book]. [Cited 8.4.2024]. https://doi.org/10.1007/978-3-031-36825-7_1.

[9] European Commission, n.d. Delivering the European Green Deal. [website]. [Cited 12.4.2024]. Available: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/delivering-european-green-deal_en.

[10] Council of the European Union & European Council, n.d. Euroopan vihreän kehityksen ohjelma. [website]. [Cited 29.4.2024]. Available: https://www.consilium.europa.eu/fi/policies/green-deal/.

[11] European Council & Council of the European Union, 2024. Pellolta pöytään. [website]. [Cited 10.8.2024]. Available: https://www.consilium.europa.eu/fi/policies/from-farm-to-fork/.

[12] European Council & Council of the European Union, 2024. Biodiversiteetti: miten EU suojelee luontoa? [website]. [Cited 10.8.2024]. Available: https://www.consilium.europa.eu/fi/policies/biodiversity/.

[13] European Commission, n.d. EU taxonomy for sustainable activities, Finance European Commission. [website] [Cited 12.4.2024]. Available: https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en.

[14] European Commission, n.d. EU Taxonomy Navigator. [website]. [Cited 12.4.2024]. Available: https://ec.europa.eu/sustainable-finance-taxonomy/.

[15] European Comission, n.d. Corporate sustainability reporting. [website]. [Cited 1.5.2024]. Available: https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en.

[16] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016. [online]. [Cited 31.7.2024]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[17] European Council & Council of the European Union, 2024. Environment Council, 17 June 2024. [website]. [Cited 3.8.2024]. Available: https://www.consilium.europa.eu/en/meetings/env/2024/06/17/.

[18] Ympäristöministeriö, 2024. EU:n ennallistamisasetus. [website]. [Cited 3.8.2024]. Available: https://ym.fi/ennallistamisasetus.

[19] Kari, M., 2022. Turvepelto-opas: RATU-hanke 2022. ProAgria Keskusten Liitto. [online]. [Cites 10.8.2024]. Available: https://www.proagria.fi/www/nettilehdet/turvepelto-opas/#/article/1/page/1.

[20] Ikävalko, K., 2024. 'Suomen 2035 hiilineutraalisuus voi jäädä vain tavoitteeksi – tutkijoiden uudessa arviossa tavoite karkaa kauas'. Yle, 22.2.2024. [Online]. [Cited 4.8.2024]. Available: https://yle.fi/a/74-20071510.

[21] Räsänen, T. A., Myllys, M., Kekkonen, H., Salo, T., Pitkänen, T., Laatikainen, M., Laine-Petäjäkangas, A., Väänänen, T., Palmu, J.-P., Kivimäki, A. & Oksanen, J., 2023. Turvepeltolohkojen märittely ja tunnistaminen : Maatalousmaiden turvetieto (MaaTu) -hankkeen raportti. Luonnonvara- ja biotalouden tutkimus 58/2023. Luonnonvarakeskus Helsinki. [online]. [Cited 15.4.2024]. Available: https://doi.org/http://urn.fi/URN:ISBN:978-952-380-720-4.

[22] Kaljonen, M., Karttunen, K. & Kortetmäki, T., 2022. Reilu ruokamurros: Polkuja kestävään ja oikeudenmukaiseen ruokajärjestelmään. Suomen ympäristökeskus. [online]. [Cited 3.8.2024]. Available at: http://hdl.handle.net/10138/349713.

[23] IBM, n.d. What is Data Science? [website]. [Cited 23.7.2024]. Available: https://www.ibm.com/topics/data-science.

[24] Kitchenham B., Charters S., 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering*, 2.3.

[25] Wohlin, C., 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE '14). Association for Computing Machinery, New York, NY, USA, Article 38, 1–10. [online]. [Cited 24.7.2024]. Available: https://doi.org/10.1145/2601248.2601268.

[26] Braun, V. & Clarke, V., 2012. *Thematic analysis*. American Psychological Association.

[27] Christiani, P., Rana, P., Räsänen, A., Pitkänen, T. P. & Tolvanen, A., 2024. Detecting Spatial Patterns of Peatland Greenhouse Gas Sinks and Sources with Geospatial Environmental and Remote Sensing Data. *Environmental management*. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1007/s00267-024-01965-7.

[28] Connolly, J., Holden, N. M., Connolly, J., Seaquist, J. W. & Ward, S. M., 2011. Detecting recent disturbance on montane blanket bogs in the wicklow mountains, Ireland using the MODIS enhanced vegetation index. *International Journal of Remote Sensing*, vol 32. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1080/01431161003698310.

[29] Ivanovs, J. Haberl, A., & Melniks, R., 2024. Modeling Geospatial Distribution of Peat Layer Thickness Using Machine Learning and Aerial Laser Scanning Data. *Land*, vol 13(4). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.3390/land13040466.

[30] DeLancey, E.R., Kariyeva, J., Bried, J.T., Hird, J.N. 2019. Large-scale probabilistic identification of boreal peatlands using Google Earth Engine, open-access satellite data, and machine learning. *PloS one*, vol 14(6). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1371/journal.pone.0218165.

[31] Tubiello, F. N., Biancalani, R., Salvatore, M., Rossi, S. & Conchedda, G., 2016. A worldwide assessment of greenhouse gas emissions from drained organic soils. *Sustainability (Switzerland)*, vol 8(4). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.3390/su8040371.

[32] Bourgeau-Chavez, L., Graham, J. A., Vander Bilt, D. J. L. & Battaglia, M. J., 2022. Assessing the broadscale effects of wildfire under extreme drought conditions to boreal peatlands. *Frontiers in Forests and Global Change*, vol 5. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.3389/ffgc.2022.965605.

[33] Kostensalo, J., Hyväluoma, J., Jauhiainen, L., Keskinen, R., Nuutinen, V., Peltonen-Sainio, P. & Heikkinen, J., 2024. Diversification of crop rotations and soil carbon balance: impact assessment based on national-scale monitoring data. *Carbon Management*, vol 15. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1080/17583004.2023.2298373.

[34] Nuthammachot N. & Stratoulias, D., 2021. Multi-criteria decision analysis for forest fire risk assessment by coupling AHP and GIS: method and case study. *Environment, Development and Sustainability*, vol 23. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1007/s10668-021-01394-0.

[35] Song, C., Choi, H.-A., Choi, E., Yang, A.-R., Lee, W.-K. & Lim, C.-H., 2024. Setting the direction of sustainable restoration projects in peatlands considering ecosystem services: Case of Jambi and Sumatra Selatan, Indonesia. *Ecological Indicators*, vol 160. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1016/j.ecolind.2024.111784.

[36] Farrell, C. A., Coleman, L., Norton, D., Kelly-Quinn, M., Obst, C., Eigenraam, M., O'donoghue, C., Kinsella, S., Smith, F., Sheehy, I. & Stout, J. C., 2021. Developing peatland ecosystem accounts to guide targets for restoration. *One Ecosystem*, vol 6. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.3897/ONEECO.6.E76838.

[37] Parthum, B., Pindilli, E. & Hogan, D. 2017. Benefits of the fire mitigation ecosystem service in The Great Dismal Swamp National Wildlife Refuge, Virginia, USA. *Journal of Environmental Management*, 203(Part 1). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1016/j.jenvman.2017.08.018.

[38] Farrell, C. A., Coleman, L., Norton, D., Kelly-Quinn, M., Kinsella, S., Obst, C., Eigenraam, M., O'Donoghue, C., Sheehy, I. & Smith, F., 2022. Applying ecosystem accounting to develop a risk register for peatlands and inform restoration targets at catchment scale: a case study from the European region. *Restoration Ecology*, vol. 30(8). [online]. [Cited 26.7.2024]. Available: https://doi.org/10.1111/rec.13632.

[39] Isoaho, A. Ikkala, L., Marttila, H., Hjort, J., Kumpula, T., Korpelainen, P. & Räsänen, A., 2023. Spatial water table level modelling with multi-sensor unmanned aerial vehicle data in boreal aapa mires. *Remote Sensing Applications: Society and Environment*, vol 32. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1016/j.rsase.2023.101059.

[40] O'Leary, D., Brown, C., Healy, M. G., Regan, S. & Daly, E., 2023. Observations of intra-peatland variability using multiple spatially coincident remotely sensed data sources and machine learning. *Geoderma*, vol 430. [online]. [Cited 25.7.2024]. Available: https://doi-org.ezproxy.cc.lut.fi/10.1016/j.geoderma.2023.116348.

[41] Bourgeau-Chavez, L., Grelik, S. L., Battaglia, M. J., Leisman, D. J., Chimner, R. A., Hribljan, J. A., Lilleskov, E. A., Draper, F. C., Zutta, B. R., Hergoualc'h, K., Bhomia, R. K. & Lähteenoja, O., 2021. Advances in Amazonian Peatland Discrimination With Multi-Temporal PALSAR Refines Estimates of Peatland Distribution, C Stocks and Deforestation. *Frontiers in Earth Science*, vol 9. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.3389/feart.2021.676748.

[42] Hugelius, G., Kuhry, P. & Tarnocai, C., 2016. Ideas and perspectives: Holocene thermokarst sediments of the Yedoma permafrost region do not increase the northern peatland carbon pool. *Biogeosciences*, vol 13(7). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.5194/bg-13-2003-2016.

[43] Leppiniemi, O., Karjalainen, O., Aalto, J., Luoto, M. & Hjort, J., 2023. Environmental spaces for palsas and peat plateaus are disappearing at a circumpolar scale. *Cryosphere*, vol 17(8). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.5194/tc-17-3157-2023.

[44] Conchedda, G. & Tubiello, F. N., 2020. Drainage of organic soils and GHG emissions: validation with country data. *Earth Syst. Sci. Data*, vol 12(4). [online]. [Cited 26.7.2024]. Available at: https://doi.org/10.5194/essd-12-3113-2020.

[45] Aich, S., McVoy, C. W., Dreschel, T. W. & Santamaria, F., 2013. Estimating soil subsidence and carbon loss in the Everglades Agricultural Area, Florida using geospatial techniques. *Agriculture, Ecosystems and Environment*, vol 171. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1016/j.agee.2013.03.017.

[46] Verdonen, M., Störmer, A., Lotsari, E., Korpelainen, P., Burkhard, B., Colpaert, A. & Kumpula, T., 2023. Permafrost degradation at two monitored palsa mires in north-west Finland. *The Cryosphere*, vol 17(5). [online]. [Cited 25.7.2024]. Available: https://doi.org/10.5194/tc-17-1803-2023.

[47] Deragon, R., Saurette, D.D., Heung, B., Caron, J., 2023. Mapping the maximum peat thickness of cultivated organic soils in the southwest plain of Montreal. *Canadian Journal of Soil Science*, vol 103. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1139/cjss-2022-0031.

[48] Negassa, W., Baum, C., Schlichting, A., Müller, J., Leinweber, P., 2019. Small-scale spatial variability of soil chemical and biochemical properties in a rewetted degraded Peatland. *Frontiers in Environmental Science*, vol **7**. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.3389/fenvs.2019.00116.

[49] Aguswan, Y., Gumiri, S., Sukarna, R.M., Permana, I, 2022. Mapping beje pond as fish source in a tropical peat swamp using Landsat 8 OLI-TIRS imagery. *IOP Conf. Ser.: Earth Environ. Sci.*, vol 1119. [online]. [Cited 25.7.2024]. Available: https://doi.org/10.1088/1755-1315/1119/1/012082.

[50] Ivankova, N. V. & Creswell, J. W., 2009. Mixed methods. *Qualitative research in applied linguistics: A practical introduction*, vol 23.

[51] Johannesson, P. & Perjons, E., 2021. *An introduction to design science*. Cham, Switzerland: Springer.

[52] Chemuturi, M., 2013. *Requirements Engineering and Management for Software Development Projects*. New York, NY: Springer New York. [E-book]. [Cited 29.7.2024]. Available: https://doi.org/10.1007/978-1-4614-5377-2.

[53] Agile Business, 2024. Chapter 10: MoSCoW Prioritisation. [website]. [Cited 5.8.2024]. Available: https://www.agilebusiness.org/dsdm-project-framework/moscow-prioririsation.html.

[54] Betz, S., Duboc, L., Penzenstadler, B., Porras, J., Chitchyan, R., Seyff, N., Venters, C. C., & Brooks, I., 2022. SusAF Workbook 6.0. [e-publication]. [Cited 20.7.2024]. Available: https://doi.org/10.5281/zenodo.7342575.

[55] Streamlit, 2024. Streamlit documentation. [website]. [Cited 31.7.2024]. Available: https://docs.streamlit.io/.

[56] Kekkonen, H., Honkanen, H., Miettinen, A., Mustonen, A., Saarnio, S., Savikko, R. Hakala, T., & Tahvola, E., 2023. Kohti ilmastoviisaampia turvepeltojen viljelykäytäntöjä. Luonnonvarakeskus Helsinki. [e-publication]. [Cited 14.4.2024]. Available: https://doi.org/http://urn.fi/URN:NBN:fi-fe2023033134220.

Appendix 1. Most and least common crops grown on Finnish peat fields

Year	Top 10 crops	Bottom 10 crops
2020	Perennial grass (dry hay, silage, soilage) Oat Fodder barley Nature management field (grass, at least 2 years) Green fallow (grass and meadow) Buffer strip (commitment since 2015) Polymorphism field, game Perennial grazing grass Spring wheat Temporarily uncultivated land	Pumpkin Savoy cabbage Brussels sprout Parsley Jerusalem artichoke Rowan Other fruit Grape Asparagus Celery
2021	Perennial grass (dry hay, silage, soilage) Fodder barley Oat Nature management field (grass, at least 2 years) Green fallow (grass and meadow) Buffer strip (commitment since 2015) Polymorphism field, game Perennial grazing grass Spring wheat Fallow	Pumpkin Celeriac Onion Annual grass seed Retiisi Rowan Other fruit Grape Grass transplant Coriander
2022	Perennial grass (dry hay, silage, soilage) Oat Fodder barley Nature management field (grass, at least 2 years) Green fallow (grass and meadow) Buffer strip (commitment since 2015) Polymorphism field, game Perennial grazing grass Spring wheat Temporarily uncultivated land	Mixed (legumes + oilseeds) Quinoa Mixed (grains + oilseeds) Spice seeds and medicinal plants Annual grass seed Energy forestry, short cycle (aspen and willow) Other lupines Natural pasture, high natural value (Åland) Asparagus Italian rye-grass seed
2023	Fodder lawn Oat Fodder barley Nature management grass	Sweet lupine Field cucumber Rowan Pumpkin

	Polymorphism crops, game	Spinach
	Green fallow (grass and meadow)	Grape
	Spring wheat	Sweet corn
	Grazing lawn	Mixed (legumes+oilseeds)
	Fallow	Kale
	Mixed (grains)	Herbs (less than 5 years, excl. dill and parsley)