Low ILUC-risk certification: Pilot report and recommendations

Ukraine, Abandoned land, February 2021

Prepared for:

European Commission

Submitted by:

Guidehouse Netherlands B.V.
Institute European Environmental Policy (IEEP)

Reference No.: 210301
ENER/C2/2018-462/LOT 2/SI2.822263
Date: 9 June 2021

Table of Contents

List of Figures

1. Pilot Introduction

This pilot was conducted to test the low ILUC-risk certification methodology for the recultivation of biomass on abandoned land, defined as "*unused land, which was used in the past for the cultivation of food and feed crops but where the cultivation of food and feed crops was stopped due to biophysical or socioeconomic constraints*" (Delegated Regulation 2019/807, Article 2(3)).

1.1 Feedstock and Geography

The 10 ha plot of land that was tested is located in Reklynec village in the west of Ukraine (see Figure 1-1) and is currently owned by a private owner/farmer. The land is currently not certified to a voluntary scheme.

The land used to be part of a state-owned farm, or "kolkhoz", that grew rye from approximately 1960-1980. After the collapse of the Soviet Union in the 1990's, the land was no longer owned by the state, and was split up into small land titles allocated to people living in the area who used to work on the kohlkoz. Often in Ukraine, these changes in ownership caused land to become abandoned, especially plots that were not very fertile.

The plot studied in this pilot was relatively sandy and degraded land, thus has only been used for some animal grazing from 1996-2021. The maximum number of cows grazing on the land reached 40 heads from 2003-2005, but on average was about 20 cows, and dropped to only 5-7 cows in recent years. It was also used for informal haymaking from 1996-2010.

During the period of abandonment (since the 1990's), vegetation naturally grew on the land. Today, the central part of the land has 6-10 year old pine trees that are approximately 2 metres tall. Satellite data confirmed that reforestation during the period of abandonment was about 15%, using the Normalized Difference Vegetation Index (NDVI).

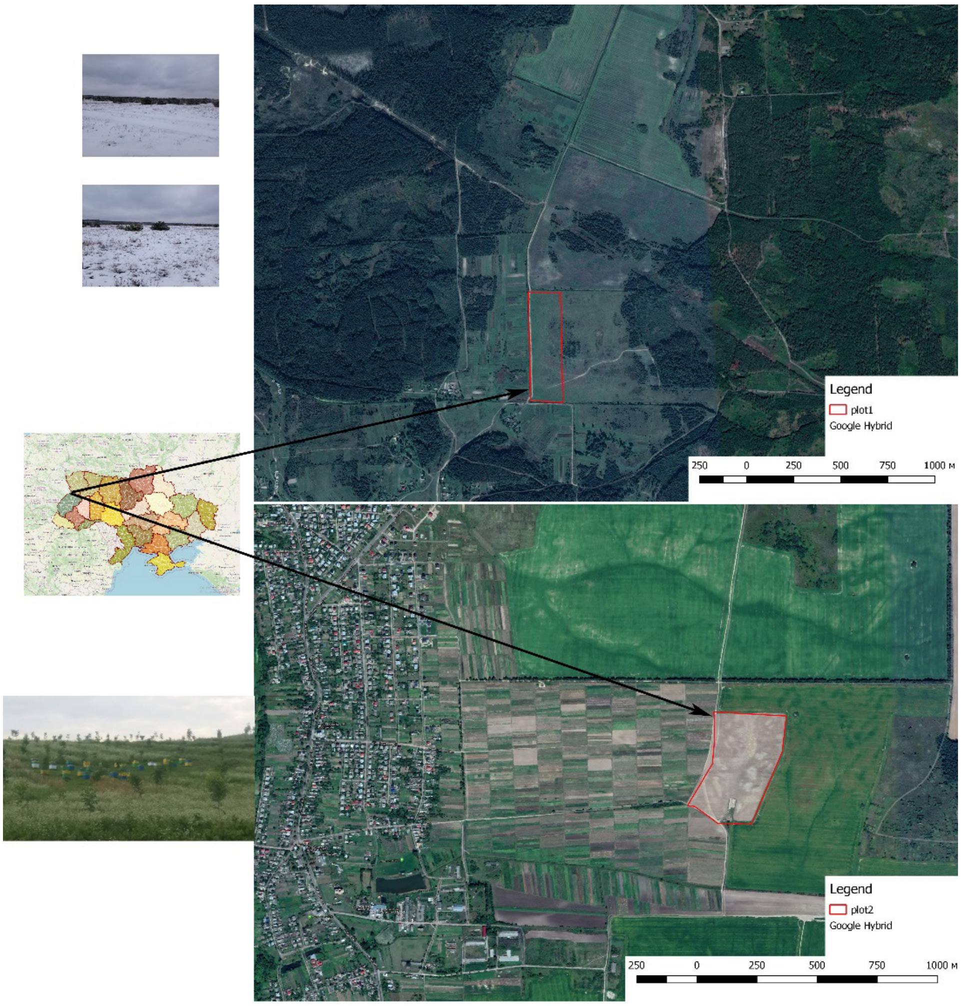


Figure 1-1. Plot of land tested in the pilot (top) located in Reklynec village in Ukraine and a neighbouring control plot (bottom) used during the satellite imaging analysis. Photos taken in 2021.

1.2 Additionality Measure(s)

The additionality measure tested is the recultivation of biomass on abandoned land that was formerly agricultural. The owner of the land has signed a land lease agreement with an agricultural company, ‘Dolyna Agro’ LLC. They intend to grow organic berries on the plot. Although this is not a biofuel feedstock, the methodology to demonstrate the status of abandonment can still be tested. The land will first grow rapeseed and soybeans to clear the land of weeds and prepare it for berry production. This is standard practice in Ukraine. These feedstocks could be used for biofuel production.

1.3 Audit

The 7-hour baseline audit was conducted onsite by Anton Opria of Control Union, an ISCC trained auditor, on **January 26, 2021**.

Oleh Chaskovskyy of the National Forestry University of Ukraine also attended the audit in person. Prior to the audit, Oleh performed a satellite imaging assessment of the pilot plot with his colleague Yuri Myklush to assess whether the plot of land met the criteria to be defined as abandoned according to Delegated Regulation 2019/807 (Annex I).

During the audit, the landowner and the co-owner of ‘Dolyna Agro’ LLC, the agricultural company who leases the land, were also physically present.

1.4 Key issues tested

In addition to testing the logic and feasibility to follow the 5 steps to identify abandoned land described in the draft low ILUC guidance, the key issues that were tested in this pilot were:

- **Satellite imaging analysis capabilities and limitations.** Satellite imaging can theoretically be used to demonstrate the type of land of a particular plot (e.g. agricultural or grassland) using the Normalized Difference Vegetation Index (NDVI). It could also potentially be able to determine the crop profile of the agricultural land to specifically demonstrate that food or feed was grown under the REDII Article 40(2) definition. This pilot tested the feasibility of this and the potential limitations of this method, such as cloud cover in images, availability of images, ability to determine crop profiles, etc (Annex I).

- **Availability of land type documentation**. Since the land was abandoned for a period of 30 years or longer, it was uncertain whether documentation to demonstrate that the land was formerly agricultural would still be available. This pilot tested the availability of such documents from government records or landowners.

- **Soil sampling for land type demonstration.** A potential complementary or substitute method to satellite imaging for demonstrating the land type is soil sampling. This pilot also explored the feasibility of using soil sampling to demonstrate that abandoned land was formerly agricultural from potential markers in the soil. This was performed independently of the audit and was a theoretical assessment (Annex II).

Note that recultivation of abandoned land is automatically considered additional, so compliance with the financial attractiveness and non-financial barrier analysis Additionality requirements in Article 5(1)(a)(i) were not tested in this pilot.

1.5 Relevant documents

As part of the audit deliverables, the following items were provided:

- Management plan (pilot company)
- Audit checklist (Control Union)
- Summary Audit Report (Control Union)
- Satellite Imaging Analysis Report: Detecting land use change, report from Oleh Chaskovskyy and Yuri Myklush of the National Forestry University of Ukraine (Annex I)
- The potential of eDNA approaches to detect historic crops in soils, feasibility study from UK Centre for Ecology and Hydrology (Annex II)

2. Findings

2.1 Long periods of abandonment and carbon stocks

According to the satellite imaging analysis, the long period of abandonment of ~30 years caused the formerly agricultural land to turn into grassland. It is unclear whether the GHG emissions associated with the direct land use change (dLUC) from grassland to agricultural land would need to be taken into consideration in the GHG savings calculation. If the land is considered as managed grassland, the cultivation of biomass on the land would not be considered a direct land use change. IPCC categorisation of managed grassland does not set a threshold for the level of grazing to be 'managed', thus grazing at any degree could be considered managed.

If the land were instead categorised as an unmanaged grassland, the dLUC associated with converting the grassland to agricultural land should be included in the GHG calculations. Assuming a yield of 1.6 tonne rapeseed per hectare as estimated during the audit, dLUC from grassland to agricultural land is approximately 109 gCO_2e/MJ. The fossil fuel comparator is 94 gCO_2e/MJ fuel, so the dLUC emissions alone exceed the GHG emissions to produce an equivalent fossil fuel. This demonstrates that the biofuel would, with certainty, not meet the GHG savings threshold in the REDII of 50-65% (relevant threshold depending on the starting operation date of the biofuel installation) if dLUC was taken into consideration.

Direct land use change emissions would be even greater if the land were categorised as having an even higher carbon stock. Despite the long period of abandonment, satellite data confirmed that reforestation during the period of abandonment was only ~15%. The plot therefore is not categorised as continuously forested areas according to Article 29(4)b, but could potentially be classified as forest land with low canopy cover 10-30%. It is not clear at this stage whether the ~15% tree cover on the plot would be removed before recultivation. If this would be the case, and the plot was (hypothetically) classified as forest land with low canopy cover 10-30%, this would result in an even greater dLUC penalty. This means that previously agricultural abandoned land which converts to forest land with low canopy cover would never be able to meet the GHG savings criteria.

It is worth noting that we found that there is not very much experience with dLUC calculations in this context in practice. Control Union indicate that they are not aware of any examples of existing voluntary scheme certification that have successfully certified biofuels with an associated direct land-use change calculation. This may be because it is simpler and lower risk for companies to just ask for biofuels with no land-use change, but also because it is very hard to meet the GHG threshold when any dLUC is included.

A potential implication of requiring dLUC to be included in GHG calculations from biomass from abandoned land is that in a situation in which a grassland was converted to a perennial crop, such as palm, there could be a positive carbon stock effect as this would increase above ground carbon stocks.

2.2 Long periods of abandonment and biodiversity

The long period of abandonment also puts into question the biodiversity that has evolved over the period of abandonment. Under Article 29(3)(d) of the REDII, biomass cultivated on highly biodiverse grassland would not be eligible due to biodiversity concerns.

In this pilot, the auditor assessed the land to be non-biodiverse grassland (see Audit checklist 3.4). This was confirmed by comparing field maps to satellite images and performing interviews with the local community. The land lease contract was also checked which stated that the land status was agricultural (legally rather than IPCC categorisation) in or before 2008. Although it was not an issue in this pilot, biodiversity is an aspect that will need to be carefully evaluated for plots that have been abandoned for a long period. In different climates, such as in tropical areas, biodiversity can also increase at higher rates as compared to Ukraine.

2.3 Satellite imaging to demonstrate food or feed crop type

Satellite imaging was readily available for the years 1986-2020 and the freely available Landsat images were used in this pilot. In general, the NDVI Index served well to demonstrate whether land was agricultural in the years studied. There were challenges for some years where only cloudy images were available and the NDVI Index could not be determined. Therefore, some years such as 2002 had to be excluded from the dataset.

The satellite imaging analysis performed by Ukrainian National Forestry University concluded that the plot of land turned from agricultural land to grassland over the period of 1986 to present day, as shown in Figure 2-1. It shows that the land was agricultural in 1986 and was grassland in 1998 as well as 2010. The challenge with the satellite imaging in the case of low ILUC biomass certification is that it is difficult to determine whether the former agricultural land specifically grew food or feed crops according to the REDII Article 40(2) definition. According to GRAS, it is technically possible to identify the crop type through satellite imaging if sufficient data is available from neighbouring plots growing the same crop. This could also be complemented with data from a country's relevant agricultural ministry on regional harvesting calendars and interviews with local community members. For this pilot however, there was the additional challenge that the land was agricultural back in the 80's, when Landsat image quality was much lower. Determining the crop profile would be more challenging using lower quality images from decades ago.

This implies that the satellite imaging might be sufficient to show if and when the land was agricultural, but may not be solely sufficient to demonstrate specifically whether food or feed crops were previously grown on the land, as required by the definition of abandoned land. The costs of such an assessment have also not been estimated in the first phase of this pilot.

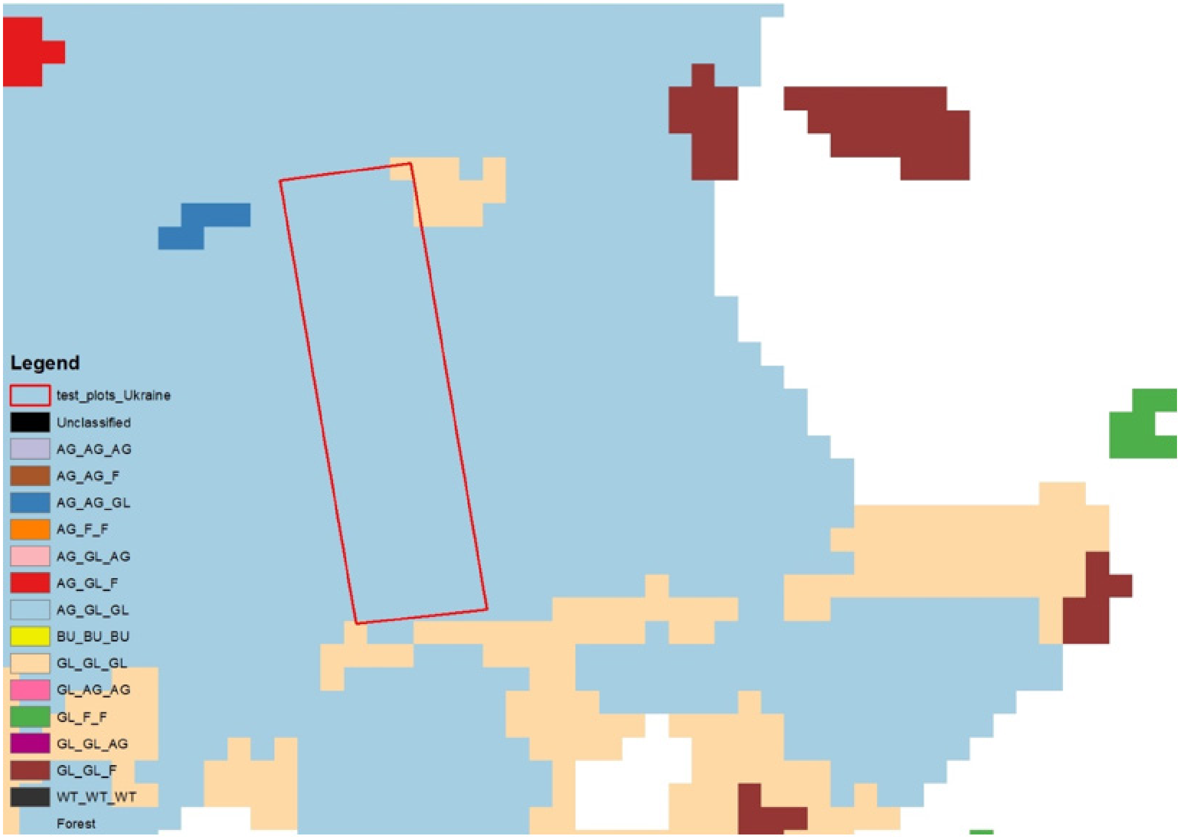


Figure 2-1. Test plot was categorised as Agricultural (AG) in 1986, Grassland (GL) in 1998, and Grassland (GL) in 2010.

2.4 Economic operators performing satellite imaging analysis

The analysis for this pilot was performed by researchers from the Ukrainian National Forestry University and cannot likely be performed in-house by economic operators. Although the Landsat images are freely available and can be analysed with free GIS software, it does require very specific GIS expertise to perform the analysis. This implies that for satellite imaging to be a viable option for economic operators to demonstrate abandoned land (without external contracting), a simple user-friendly tool would need to be made available. Alternatively, satellite imaging analysis would need to be contracted, and would add to the cost of certification.

2.5 Grazing during the period of abandonment

According to GRAS, extensive and intensive grazing can be identified with satellite imaging. However, small-scale grazing of animals was indicated as near impossible to demonstrate through satellite imaging. It would require a large amount of data and would still need to be confirmed with groundtruthing (costs were not assessed as part of this pilot). The absence of grazing thus cannot be confirmed with certainty through satellite imaging.

In this pilot, the evidence ultimately used for grazing was through the auditor conducting three interviews with local community members aged 76-82 years who had lived in the village since before the period of abandonment. The interviewees indicated that there was some grazing during the period of abandonment. They indicated that the maximum number of cows grazing on the land reached 40 heads from 2003-2005, but on average was about 20 cows, and dropped to only 5-7 cows in recent years. Since there is no definition of ‘substantial grazing’ provided in the legislation, it is difficult to assess whether this number of livestock and frequency is considered as ‘substantial’. Currently it would be up to the discretion of the auditor whether interviews provide sufficient certainty.

2.6 Local archive availability

The period of abandonment of the plot of land was nearly 30 years, far exceeding the minimum of 5 years in the legislation. Ukraine kept thorough agricultural documentation from the Soviet period, but it was uncertain whether local archives would still have these documents available from this period.

In this pilot, we learned that in Ukraine, local administration only keeps for documents for 25 years, thus there was no documentation to prove or disprove whether the land was agricultural and growing rye from 1960-1990. Satellite images were however available from 1986 onwards, thus could demonstrate that the land was agricultural from 1986-1990. The interviews performed with locals who lived in the area during this period also complemented and confirmed this.

2.7 Estimated future additional biomass

The economic operator estimated the future yield of rapeseed to be 1.6 t/ha, based on yields achieved in other local farms. In the context of recultivation of abandoned land, this estimate is not needed for the financial attractiveness test.

The main use of this estimate in this context is for an auditor to judge whether the achieved additional biomass after recultivation of the land is reasonable and in an expected range. This avoids any possibility of fraudulent claims of additional biomass by a farm claiming material not grown on that plot of land to be additional.

The estimate of 1.6 t/ha is relatively low but still within the range that would be expected in this context, also given that the farmers here are often growing rapeseed with the intention of clearing the land of weeds ready for berry farming, rather than aiming to optimise rapeseed yields.

2.8 Soil sampling to demonstrate formerly agricultural land

Soil sampling is not advisable to be the sole method to demonstrate that land was previously agricultural and used to grow food or feed crops, although it could potentially be used to complement satellite imaging analysis or local interviews. In a report performed by the UK Centre for Ecology and Hydrology for this project, it was concluded that environmental DNA (eDNA) can be extracted from soil samples to determine the crop species that used to be grown in the soil. This is done through the highly developed technology of DNA sequencing, and public DNA sequence database repositories are available for common crop species such as sugars and oilseeds.

The difficulty is determining the historical timescale of detection, i.e. concluding with certainty that a crop was grown 7 years ago versus only 4 years ago. Although DNA can persist in the environment for many years, the detection and persistence of eDNA in soils is influenced by many factors including soil environmental variables such as pH, UV radiation, alongside biological factors such as microbial degradative activities.

Soil sampling was not performed during this pilot, but as an example, if eDNA from rye was detected from a soil sample, it could not be said with certainty that this crop was only grown in the 1980's and not during the period of abandonment. Therefore, it can be used as one, but not the only tool to demonstrate that land as formerly agricultural. The cost of sequencing eDNA is relatively low, between 50-100 EUR per sample, however the additional costs to interpret and report on the sequencing could potentially significantly increase costs.

3. Recommendations for low ILUC-risk methodology

Based on the main findings from the pilot, we suggest the following recommendations for the low ILUC-risk methodology:

Carbon stocks and abandonment. The definition of abandoned land in the REDII requires land to have been abandoned for a minimum of 5 years. The period of time it takes land to convert from agricultural land to grassland or even forest will depend, amongst others, on the climate conditions and the use of the land during the period of abandonment. After 5 years it might be possible that the land could still be classified as agricultural and therefore no emissions from direct land use change would need to be taken into account. However, after longer periods, the land status might have changed (e.g. it might be considered to be a grassland or forest land). In that case, it would be challenging for the land to be converted to agriculture and comply with the REDII GHG saving criteria for biofuels, once dLUC is taken into account. That will depend on a case-by-case basis. The methodology should further detail when dLUC should be taken into consideration as this will ultimately determine if biofuels from this land can comply with the GHG criteria.

Biodiversity and abandonment. Long periods of abandonment also put into question how the biodiversity of the land may have increased during the period of abandonment. Auditors should diligently assess the biodiversity of the land before it has been recultivated.

A clearer definition is needed for substantial grazing. With the current methodology, the lack of a clear definition of “substantial” leaves it to the discretion of the auditor whether the land meets the definition of abandoned land. A threshold such as number of livestock per hectare, or clearer guidance to auditors which criteria can be used to assess whether grazing is substantial could be considered. However, such a threshold is likely to vary quite broadly per region and per type of animal so this may be challenging in practice. In addition, from an ILUC perspective, it could also be an important factor to consider whether there is other land available in the proximity for animals to graze, if displaced. These type of considerations for auditors to examine could help guide them to make a decision.

If satellite imaging is to be used as a tool to demonstrate abandoned land, external experts or a centralised platform is needed. It has been demonstrated that satellite imaging could be a powerful tool for demonstrating abandoned land in the absence of other documentation or interviews. However, this assessment cannot be done by economic operators unless they have in-house GIS expertise. If satellite imaging is to be relied on as a tool to demonstrate abandoned land, a centralised platform for economic operators to enter the geographical coordinated of their plot and receive a land history status assessment could help facilitate this. Some previous studies[1] have suggested that such a platform would cost in the order of 160-250 kEUR.

[1] Ecofys (2016). Methodologies for the identification and certification of Low ILUC risk biofuels. https://ec.europa.eu/energy/sites/ener/files/documents/ecofys_methodologies_for_low_iluc_risk_biofuels_for_publication.pdf

4. Next phase of the pilot

In the next phase of the pilot, we recommend to either i) further test challenges that were faced during this pilot on the same plot of land, or ii) test new aspects of the methodology on the same plot, or iii) test new aspects on a new plot of land.

The aspects that were not fully solved in this first baseline audit that could be further explored are:

- The issue of substantial grazing and defining more concrete criteria for auditors
- Determining the crop profile of a food or feed crop from satellite imaging

In a future additionality audit, the aspects that could be tested are:

- The actual yield of low ILUC biomass (additional biomass) achieved from the converted land
- Cost of converting land

However, these are both contingent on the timeline of when the agricultural company will begin to grow the rapeseed. If this does not align with the timeline of Phase 2 of this project, then this will not be possible.

A different plot of land could be studied to test the following aspects:

- Demonstrating land that is severely degraded
- Financial attractiveness and non-financial barrier analysis for ‘other unused land’
- Additional biomass for unused land – especially if the land is growing a crop that is not a food/feed crop, e.g. coffee

A new plot of abandoned land for Phase 2 could be found through the network of stakeholders. Alternatively, a study could be done to quantify the land that would meet the abandoned land criteria in a given area and identify potentially interesting plots to study.

guidehouse.com

Satellite Imaging Analysis Report

Detecting land use change

Submitted by: Oleh Chaskovskyy and Yuriy Myklush

National Forestry University of Ukraine

Table of Contents

Definitions:

Unmanaged grassland – part of an abandoned land (no human activities, possibly used for the grazing of the small amount of livestock) where the vegetation is dominated by grasses.

Managed grassland – part of a land managed by humans for haymaking

Introduction

Description of Plots of Land

Local stakeholders who own two plots (shown in sent shape file) are interested in cooperation and ILUC certification. Although these plots are located close to the EU border, in particular to Poland. A lot of products from this area are exported to the EU. These plots are characterized by a set-up relevant for the Polissya region (which is a big part of Ukraine) and poor soil conditions that consequently stipulated abandonment of agriculture areas after Soviet Union collapse.

In addition, they were unmanaged until recently. The history of their governance is unknown, therefore deep analysis based on satellite images is required.
Likewise, maps based on previous researches created for this region by scientists from Humboldt University in Berlin were used for selection. Plots that were classified as abandoned lands were chosen (Fig. 1).

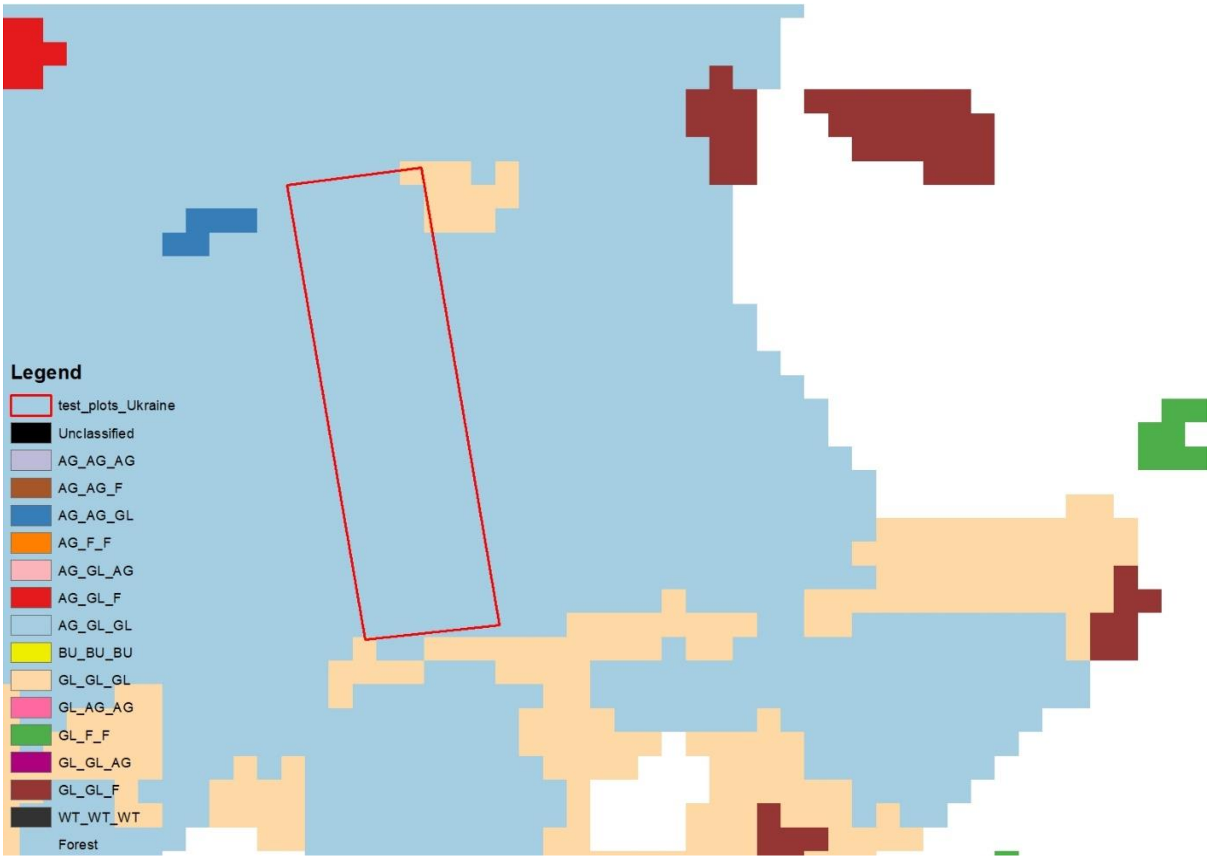


Figure 1. *Test plot #1 classification as AG_GL_GL (agriculture_grassland_grassland) for the period from 1986 to 2010*

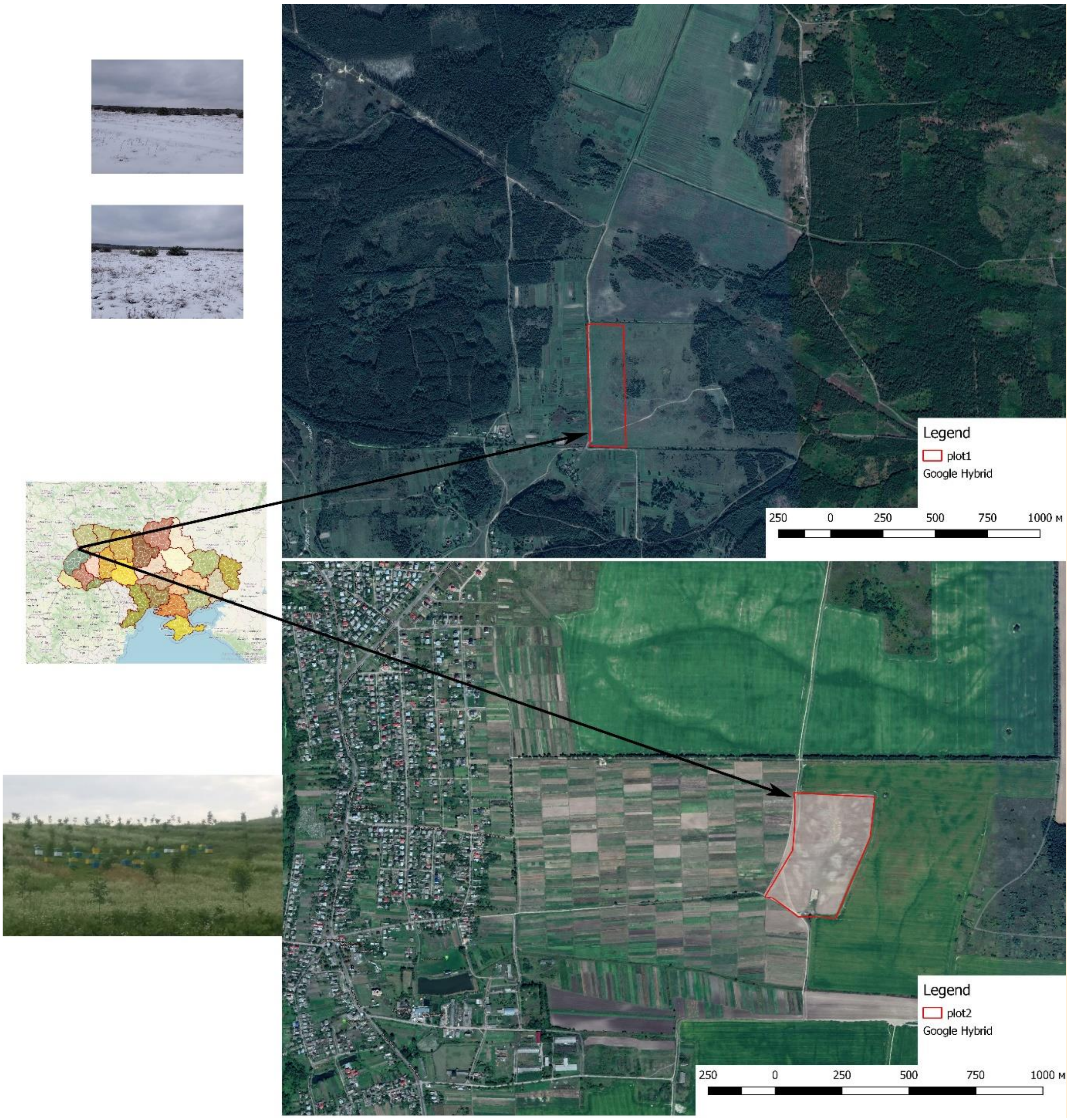


Figure 4. *Pictures of land and geographical coordinates for investigated plots*

The coordinates of the centroids of the plots are presented in Table 1.

Table 1

Location and area of test (plot1 and plot2) and control plots (control1 and control2)

Name	xcoord	ycoord	area, ha
plot1	24,25353	50,24701	10
plot2	24,30685	50,45372	14
control1	24,26107	50,26085	10
control2	24,3119	50,45296	14

Since both plots from the 1990s until recently were abandoned, we selected nearby control plots of the same area as the studied plots (Fig. 2).

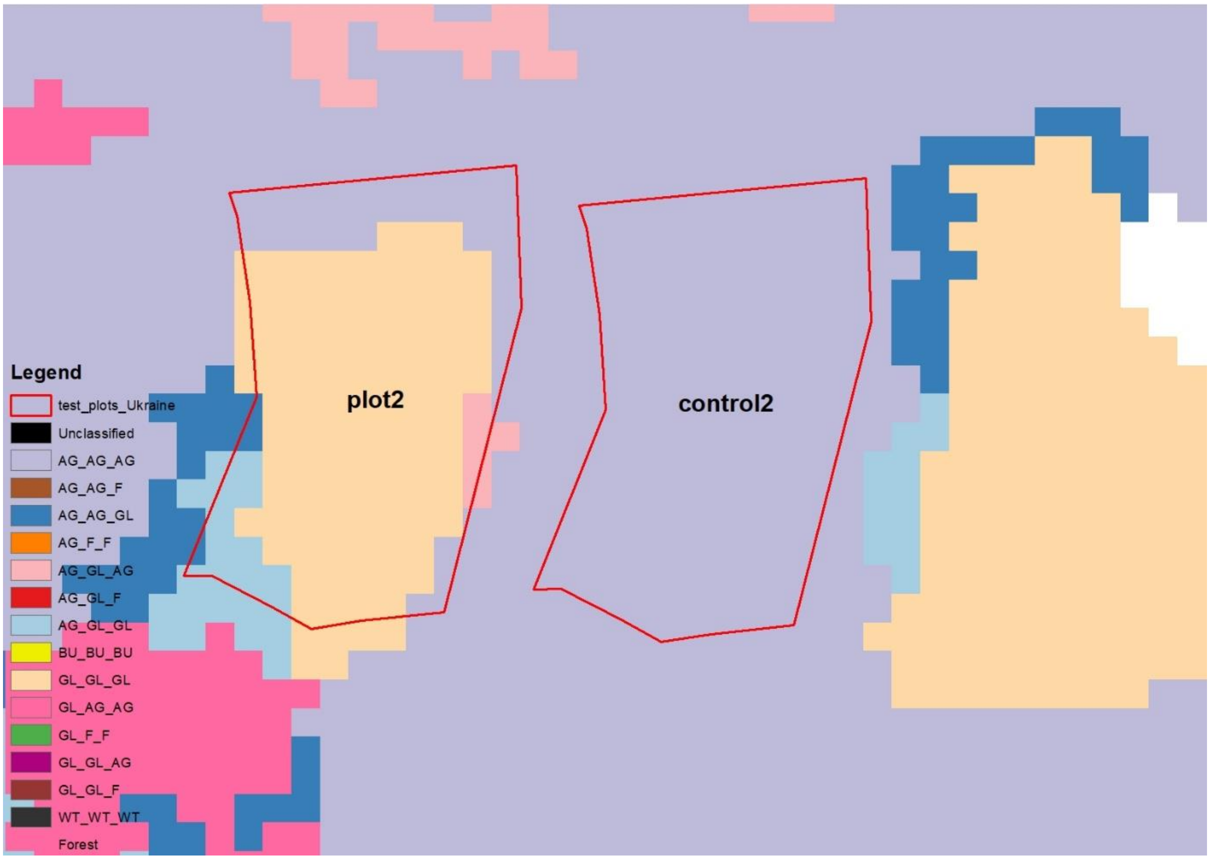


Figure 2. *Plot #2 and control plot with same area*

Plot #2 is mainly defined as a grassland (GL_GL_GL), and the control plot, which is equal in area and similar in shape, is decoded as managed field (AG_AG_AG). This control area will be used as a model for determining the constantly managed field (AG). Regarding plot #1, a control plot was determined on the basis of visual analysis (because according to the available maps until 2010 there were no plots of managed field (AG) nearby).

Table 1a provides an overview of the status of plot#1.

Table 1a.
Status of the plot#1 from 1974 to 2020

Time periods	Status of plot#1
1974 - 1994	**Managed grassland (used for haymaking**). For the period from 1986 to 1994 status confirmed by decoded satellite imaging; earlier confirmed by local interviewers
1994 - 2020	***Unmanaged grassland*** *(an abandoned land without*

	human activities; unintentionally could be used for the grazing of the small amount of livestock). Status confirmed by local interviewers

Three sources of information have been used to detect land use change:

1) Raw images for visual interpretation*
2) Processed images*
2a) Preprocessing process includes DOS1 correction, which is natively included in plugin called Semi-Automatic Classification Plugin (https://plugins.qgis.org/plugins/SemiAutomaticClassificationPlugin/) for QGIS. DOS "Dark Object subtraction" is a method for atmospheric correction. DOS is one of the most commonly used atmospheric correction methods (Zhang & Wang 2010).
2b) Calculating NDVI values for each pixel of both plots of interest for all images based on preprocessed images by using Zonal Statistics Plugin and native QGIS function "Raster calculator". Explanation of NDVI is provided below. Taking into account that NDVI values are calculated for each pixel (not for a plot), four statistical values for each plot were determined, namely minimum value, maximum value, mean value, majority.

The images of satellite Landsat were analyzed for the period from 1986 up to 2020. All available relevant images were preprocessed and post-processed in QGIS (special open-source software for GIS, described below).

3) Data from interviews with the local representatives (farm workers, stakeholders, authorities)

Methodology

All images were downloaded from https://earthexplorer.usgs.gov/. U.S. Geological Survey provides the ability to download images for free after successful registration.

Open source software QGIS (*https://qgis.org/*)(detailed description on page 2 of this report)

Normalized Difference Vegetation Index (NDVI) is used to determine the density of green on an individual pixel. When sunlight strikes objects, certain wavelengths of this spectrum are absorbed and other wavelengths are reflected. The pigment in plant leaves, chlorophyll, strongly absorbs visible light (from 0.4 to 0.7 µm) for use in photosynthesis. The cell structure of the leaves, on the other hand, strongly reflects near-infrared light (from 0.7 to 1.1 µm). The more leaves a plant has, the more these wavelengths of light are affected, respectively (NASA 2020).

Nearly all satellite Vegetation Indices employ this difference formula to quantify the density of plant growth on the Earth — near-infrared radiation minus visible radiation divided by near-infrared radiation plus visible radiation (Fig. 5). The result of this formula is called the Normalized Difference Vegetation Index (NDVI). Written mathematically, the formula is (NASA 2020):

$$NDVI = (NIR — VIS)/(NIR + VIS)$$

Figure 5. *Meaning of NDVI values (picture taken from https://eos.com/blog/ndvi-faq-all-you-need-to-know-about-ndvi/)*

Calculations of NDVI for a given pixel always result in a number that ranges from minus one (-1) to plus one (+1); however, no green leaves give a value close to zero. A zero means no vegetation and close to +1 (0.8 - 0.9) indicates the highest possible density of green leaves (NASA 2020).

According to the data from interviews, Plot #1 had been used to grow crops until the 1970s, later on soil drainage was performed and after that land use type changed to managed grassland. This type of land use existed until kolkhozes collapse in 1994. From 1994 until nowadays plot #1 has had the status of grazing land (the amount of livestock approximately 20 animals and their influence on the plot surface has been insignificant). Therefore, we can check the hypothesis about changes in NDVI values for managed grassland and since 1994 for unmanaged grassland.

Nearby plot classified as agriculture (AG) was chosen for comparison based on Landsat-8 image (Fig. 6)

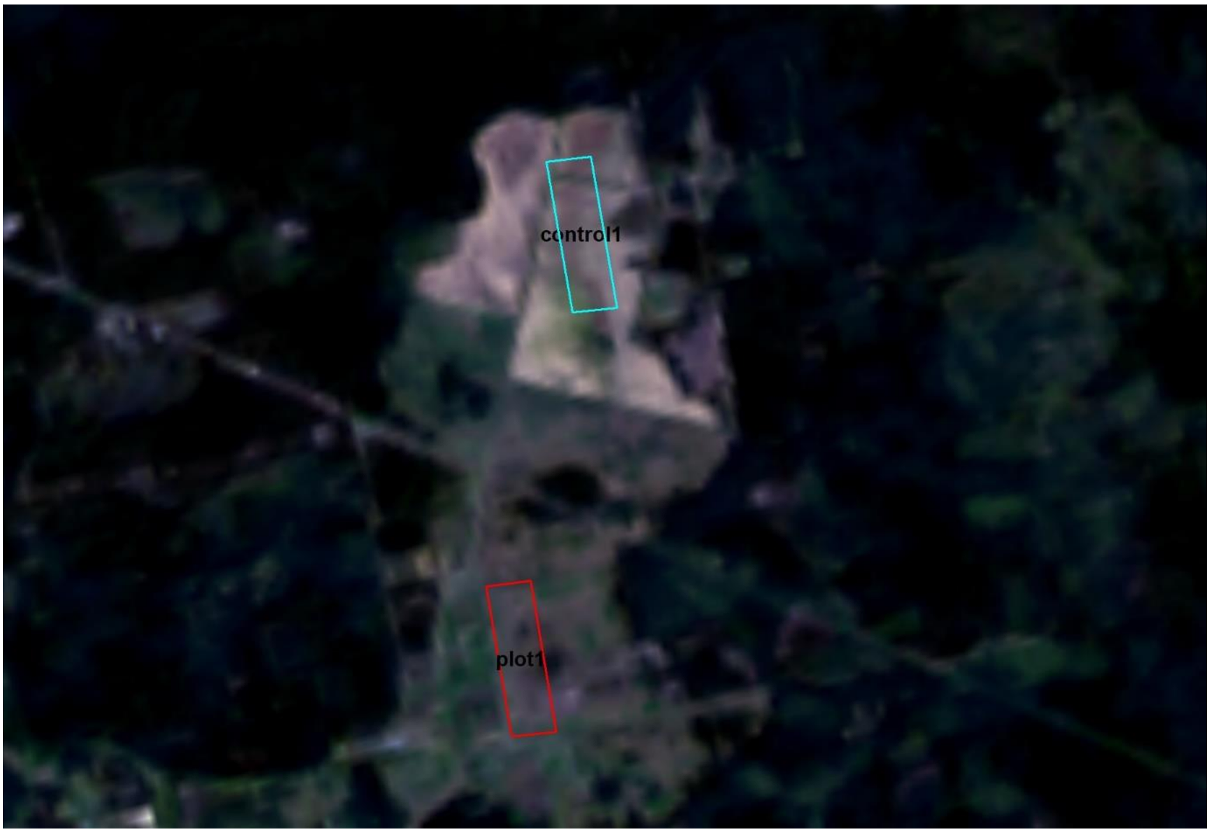


Figure 6. *Plot classified as agriculture (AG) used for comparison based on Landsat-8 image as of August 2019*

According to our research, to use visual interpretation and NDVI values of processed images for detecting land use change at least 3 images of the same year are required. The change in data usually can be observed during April-early May (sowing), June-July (vegetation), August-September (harvesting) which can be explained by regular crop rotation. According to our hypothesis short crop rotation period should be observed on plot #1.

Categorization of Land

Abandonment was categorized in case insignificant or no change detected on a plot between three values of NDVI index within vegetation season (from April to August) per annum.
Managed grassland was categorized in case significant change detected on a plot between three values of NDVI index within vegetation season per annum (can be explained by April-early May (sowing), June-July (vegetation), August-September (harvesting).

Cloudy Images

All available images (even with high cloud cover) were downloaded and analyzed. In total we analyzed 99 images of two scenes (185_25 and 185_26) (fig. 7). Images where plot territories of interest were covered by clouds (or any other obstacle) were skipped. That is the case for the year 2002.

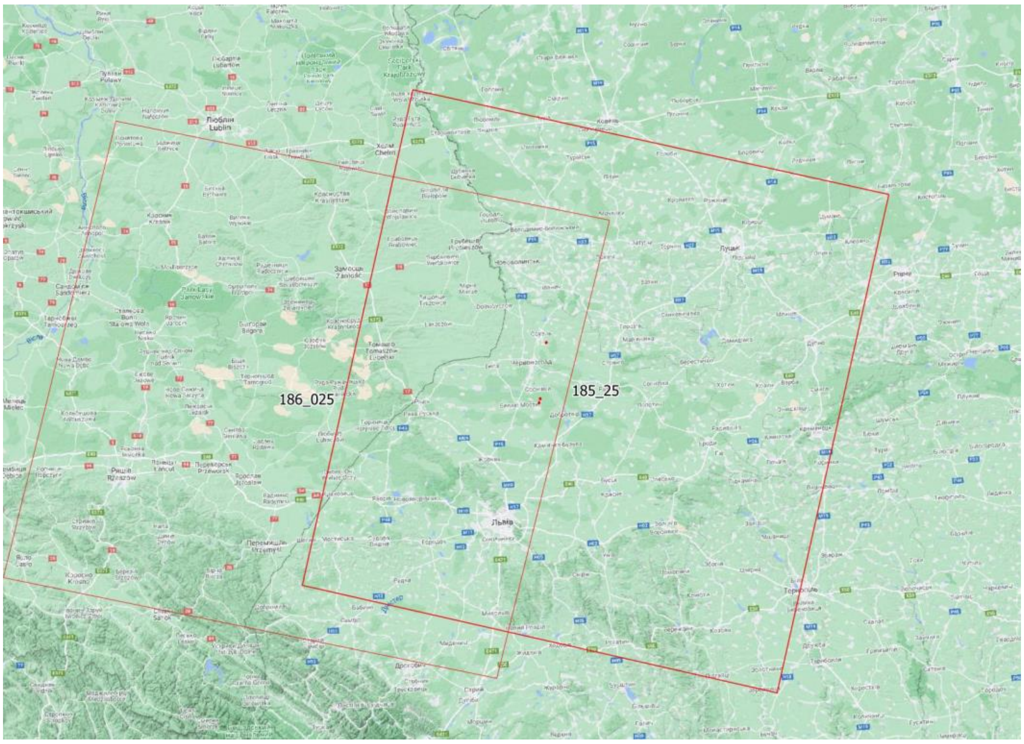


Figure 7. *Scenes (path_row) of the test plots (regions)*

Results

Landsat images are freely available and their resolution allows detecting land use changes.

To determine the changes in land use it is possible to use a visual interpretation of the studied areas as well as changes in spectral characteristics of satellite images during the season associated with plowing and harvesting. Such a visual analysis gives possibility to determine field boundaries in case farming activities took place.

The values of NDVI are calculated for the period from 1986-2019, analyzed in details within 2008-2019. For each year, images for the growing (vegetation) season (spring-summer-autumn) are selected. Moreover, for the period 2008-2011 images of Landsat 4-5 had been used, for 2012 only Landsat-7 images are available and images of Landsat-8 have been used since 2013.

The results of the calculation of NDVI and their mode values calculated for each plot are presented in the graph (Fig.8).

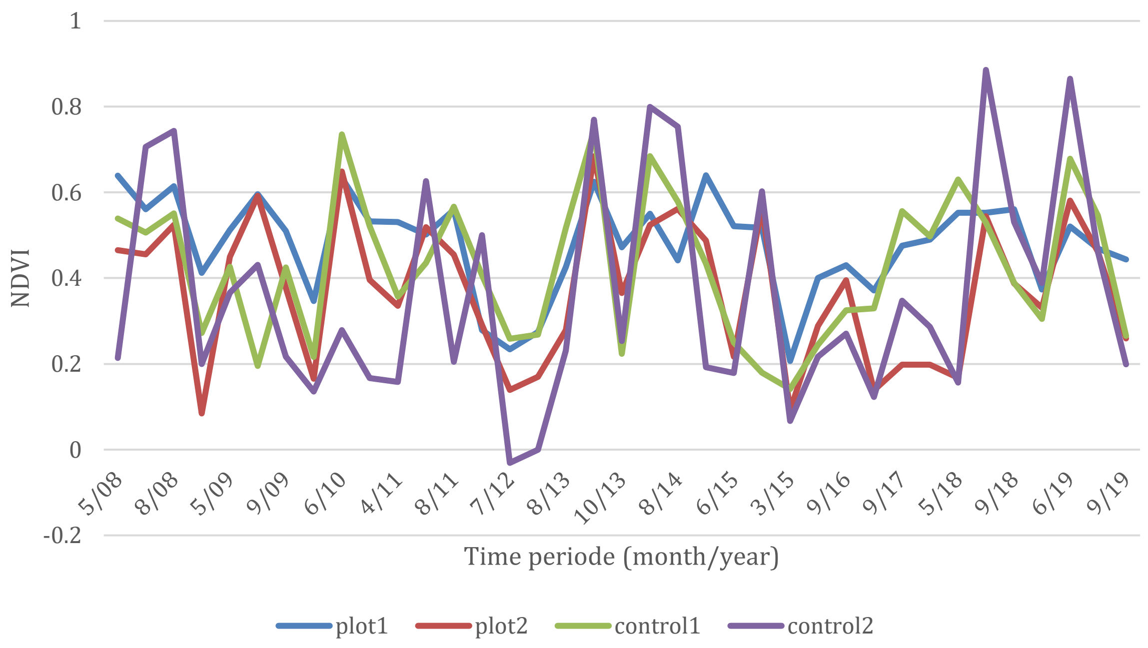


Figure 8. *NDVI values for the test and control plots for different time periods*

The figure shows that the NDVI values of the used control plots have a greater variation during investigated period. This is especially true for the control plot #2 that had been constantly used for agricultural production. The values of NDVI on this plot vary from 0 to 0.9 that can be explained by the reflection of soil surface during plowing and dense vegetation during the growing season. For control plot 1, the difference is significant only after 2010, when the plot began to be used. Moreover, the category of hayfield has less variation and does not reach values of 0 but only values of 0.2 which is close to the values of liquefied vegetation

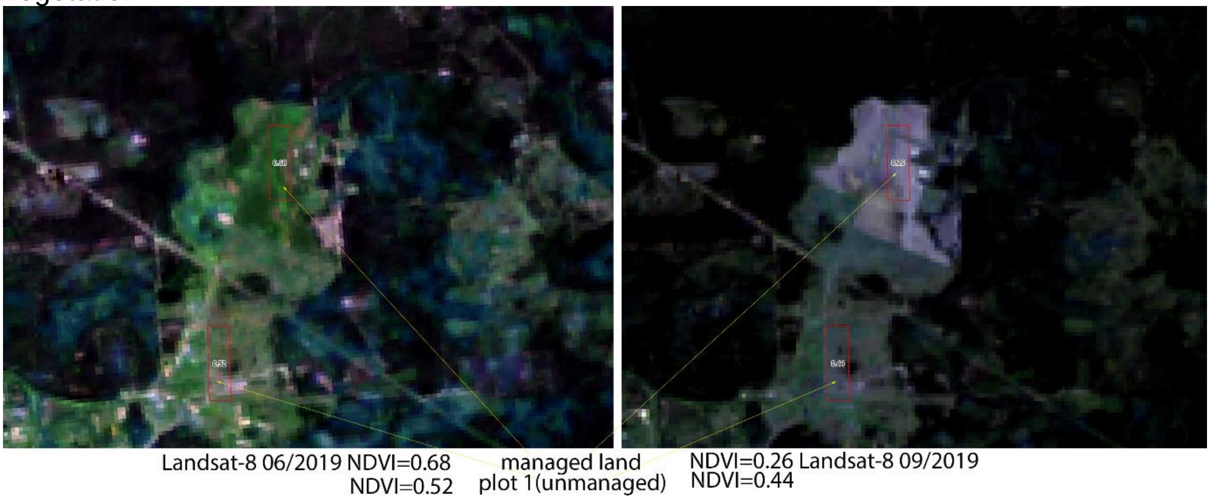


Figure 9. *Nearby plot classified as agriculture (AG) was chosen for comparison based on Landsat image.*

The values of NDVI of the plot classified as agriculture (AG) changes from 0.68 to 0.26 and of abonnement plot (unmanaged) – from 0.52 to 0.44.

Limitations

Crop Profile

The type of crop grown cannot be determined from satellite images based only on NDVI values due to the fact that values of NDVI of different types of crops are extremely insignificant (for some types of crop they even can be equal). Therefore, for such analysis we need to know crop rotation period for each possibly grown crop profile and to have at least three high resolution images taken within the vegetation period of those crop profiles.

Grazing

Satellite imaging can prove grazing of livestock only in case significant influence by animals happens on a grassland when all the vegetation is destroyed or eaten (in case of big farms that is not relevant for our area of interest). It cannot be proven with satellite imaging that grazing of livestock did take place for small farms or single cattle due to the fact that values of NDVI, especially mean values and majorities, will differ slightly and consequently it might be hard to prove that exactly grazing of livestock stipulated that difference.

Uncertainty

Landsat satellite imagery is not always available or cloud-free for each required time period. It is necessary to review all the images to determine whether they are suitable for analysis of specific areas. In our case it was not possible to select images for the year 2002. In the case research is required for a long period, only images from different satellite systems are available (Landsat4-8). For the year 2012 only Landsat-7 images with defects (stripes) are available and only one of those image was found and used without defects, as the study area is located in the center of the image.

Conclusions

Landsat satellite images have been freely available since 1986. Their visual interpretation makes it possible to determine the fields boundaries and their management, which can be used as a criterion for assessing the abandonment of the plot. Also, the change in NDVI values is a reliable tool for such analysis. A large scale of seasonal variation from 0 to 0.9 indicates a managed field, and a smaller one: from 0.2 to 0.7 – about haymaking. The increase in the variation of NDVI values indicates the beginning of land use that had been previously abandoned.

References

1. Normalized Difference Vegetation Index (NDVI). NASA 2020. available on https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php
2. Zhaoming Zhang, Guojin He & Xiaoqin Wang (2010) A practical DOS model-based atmospheric correction algorithm, International Journal of Remote Sensing, 31:11, 2837-2852, DOI: 10.1080/01431160903124682

The potential of eDNA approaches to detect historic crops in soils

Can DNA techniques be used to confirm that particular crops were grown on a field many years ago?

Rob Griffiths[1], David Mottershead[2], Tim Goodall[1], Dan Read[1] & Chris Bell[1]

[1] UK Centre for Ecology & Hydrology, [2] Institute for European Environmental Policy

Client Ref: Project number 210301

Issue Number 0.9

Date 31/03/2021

Series	--
Title	The potential of eDNA approaches to detect historic crops in soils
Client	Guidehouse Netherlands B.V.
Client reference	Guidehouse/Project number: 210301
UKCEH Reference	WD07848
Confidentiality, copyright and reproduction	
UKCEH contact details	UK Centre for Ecology & Hydrology (UKCEH) Environment Centre Wales, Deiniol Road, Bangor Gwynedd, LL57 2UW 01248 374500
Corresponding author	Rob Griffiths, UKCEH rig@ceh.ac.uk
Authors	Rob Griffiths[1] , David Mottershead[2], Tim Goodall[1], Dan Read[1] & Chris Bell[1] [1] UK Centre for Ecology & Hydrology, [2] IEEP
Contributing authors & reviewers	Our thanks to the creators of https://undraw.co for images used.
How to cite (long)	Griffiths, R., Mottershead, D., Goodall, T., Read, D. & Bell, C. (2021) The potential of eDNA approaches to detect historic crops in soils. Report to Navigant Netherlands/Guidehouse (Project 210301)(UK Centre for Ecology & Hydrology Project 07548)
How to cite (short)	Griffiths, R. et al. (2021) The potential of eDNA approaches to detect historic crops in soils. Report to Navigant Netherlands/Guidehouse (210301)(UKCEH 07548)
Approved by	
Issue Date	

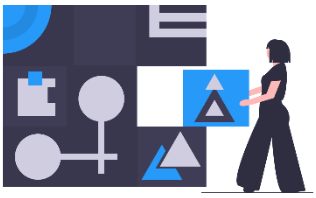
Contents

Abbreviations Used in this Report

DNA	Deoxyribonucleic acid
eDNA	Environmental DNA
IEEP	Institute for European Environmental Policy
ILUC	Indirect Land Use Change
in silico	Research conducted or produced by means of computer modelling or computer simulation
ITS	Internal transcribed spacer [marker/spacer DNA]
PCR	Polymerase Chain Reaction
UKCEH	UK Centre for Ecology & Hydrology

1 Introduction

1.1 Context



Navigant Netherlands B.V. and the Institute for European Environmental Policy (IEEP) have commissioned the UK Centre for Ecology & Hydrology (UKCEH) to create this report on the utility of "eDNA" methodologies for soil testing in respect to certification of "low ILUC risk" energy feedstocks (industrial or transport related biofuels).

ILUC refers to "Indirect Land Use Change" and relates to the situations where agricultural land used traditionally for food and feed is converted to biofuel production. In terms of land use change accounting, this can result in intensive land use expansion to maintain a similar level of food and feed production (e.g. intensification of forests/peatlands), thereby offsetting any benefits from biofuel production.

One way in which the EU Renewable Energy Directive permits such feedstocks to be certified is where they are used on abandoned land on which food or feed crops were grown in the past, but such production was abandoned for a period of at least five years. The requirement in the legislation is to prove that any sort of food or feed crop was grown. This definition broadly refers to oilseeds (rapeseed, soybean, sunflower, palm, etc.), starches and sugars (e.g. maize, potato, sugar beets/cane, wheat, rye, barley, etc.). There are various ways of demonstrating that such crops grew in the past, notably by checking archives in those countries with accurate records. However, other approaches are required where such evidence is lacking, and there is interest in establishing whether soil testing can meet this requirement. One such promising methodology is the use of soil DNA to test for historic residues indicative of these crops.

1.2 Scope of this work

A short report on the feasibility of using eDNA from soil samples to establish that food and feed crops have been grown in the past. How far into the past can a soil test see, and with what accuracy as regards crop types and individual years?

2 The potential of soil eDNA for plant species detection

All living organisms possess a unique DNA code, which is now accessible for research and diagnostic purposes, in large part due to recent advances in high throughput DNA sequencing technologies (Deiner et al., 2017). The use of DNA sequences as a taxonomic marker for plant and other species is well established, and a variety of "marker gene" diagnostic regions have been proposed based on sequencing of collected plant specimens.

The recent advancements permitted by the new sequencing technologies termed Environmental DNA (eDNA) analyses, relates to the analyses of *DNA extracted from the environment* rather than an isolated species specimen (for reviews and a discussion of the terminologies used see Thomsen & Willersley (2015), Ruppert et al. (2019) and Pawlowski et al (2020). Following the extraction of DNA from an environmental matrix such as soil, sediment or water, a pool of genomic DNA is obtained which is a mixed sample comprising of DNA from the multitude of organisms present in the sample. A marker gene of interest can then be amplified with broad specificity PCR primers, and then specific species identified and partially-quantified through the use of high throughput sequencing.

eDNA approaches are well developed and applied for assessments of microbial communities, but recently there is increasing interest in applying them to the assessment of larger organisms such as plants and invertebrates. Importantly, since remnant DNA from an organisms can be present in residues which can persist in different environments, the methods can be applied to small volumes of sample, for instance a crumb of soil. There is therefore much interest in using eDNA for plant species monitoring, and it offers the potential for simplifying the field effort required, as well as enabling the processing of vast numbers of samples with high taxonomic resolution, down to species or even sub-species level. eDNA analysis has provided insight to many, otherwise difficult to monitor, environments and is assisting in estimates of biodiversity and distribution of both micro and more recently macro organisms (Ruppert et al., 2019; Thomsen & Willerslev, 2015).

The accuracy of eDNA for plant community analysis from soil is relatively novel and untested at scale, though Fahner et al. (2016) assessed a suite of plant taxonomic markers on 35 boreal wetland soils. In general there are only a handful of papers addressing plant eDNA detection in soils specifically, though there are historical papers addressing plant DNA persistence in soils (from earlier research aiming to look at persistence of genetically modified elements). The application of these methods for detection of *historic* plant residues for policy purposes, is highly novel and largely untested. There is therefore no direct scientific literature available for reporting on efficacy, and so the following report draws on the broader limited literature discussing plant eDNA detection from soils.

2.1 Species accuracy of detection

For it to be of general application in the identification of abandoned land, any eDNA approach used needs to be capable of detecting the major crop species relevant to the policy. However even a test capable of identifying only a few species, or even a specific plant, could be of use in situations where the former presence of those species is suspected. Whether commonly used 'off the shelf' plant diagnostic assays will detect all species of interest depends on two factors:

i) The completeness of the reference database - have all species of interest been sequenced and publically archived?
ii) Primer specificity - do the primer sequences used to amplify the metabarcoding gene region of interest adequately amplify all the species of interest?

A study published by Fahner et al. (2016) quantified the availability of sequence information for plant species held in public DNA sequence database repositories, and although it focussed on wetland plant species, the majority were found in the reference database and the "ITS" marker gene region possessed best coverage.

This is confirmed in our own work on GB wide soils with standard, published ITS assays where we detect most common crop species (sugars and oilseeds in the UK). Therefore the use of current standard molecular assays will likely be suitable for the specific policy goals to identify general crop species (which are likely to have been sequenced); and furthermore, accurate cultivar discrimination (e.g. specific strains of wheat) is not a primary objective (further research would be required if the objective was for strain level diagnostic assays). However, as a matter of rigour, before implementing eDNA assays it would be advisable to check database coverage, e.g. http://its2.bioapps.biozentrum.uni-wuerzburg.de/ or using the methods described in Fahner et al., (2016), to check if reference DNA sequences from most species representing the broad categories of crops are available, as well as lab validation of the assays on reference soils known to contain the species.

In the unlikely event that there are certain species not covered, some research activity may be required to enhance PCR assay coverage, though this would not likely be problematic for specialist labs carrying out such work.

2.2 Historical timescales of detection

The critical issue in using eDNA approaches to detect past instances of cultivation is the length of time that DNA persists in the environment. It is well established that DNA can persist in the environment for many years; for example, DNA has been successfully amplified from lake sediments that are more than 10,000 years old (Crump et al., 2021).

In our own work on the *UKCEH GB: Countryside Survey*[1] we find large amounts of plant diversity in eDNA inventories from arable soils, which is at odds with the monocultures of crop often recorded by the field surveyor and emphasises that contemporary soil DNA contains relics of past species present. The detection and persistence of eDNA in soils is influenced by many factors including soil environmental variables such as pH, UV radiation, alongside biological factors such as microbial degradative activities (reviewed in Levy-Booth et al. (2007). Of particular importance, the review identified that the soil mineralogy (sand, silt, clay content and type) has strong impacts on how much DNA gets bound to the soil matrix, though the relevance for persistence is unknown.

Only two recent studies have specifically investigated historic plant DNA detection in soils. Yoccoz et al. (2012) assessed formerly cultivated plots abandoned between 1810 and 1986 in temperate environments and found that the number of historic crop DNA sequences identified varied over time since last cultivated. In general, a negative relationship was observed between the number of historic sequences detected and years since abandonment, with no historic DNA detected after ~50 years.

---

[1] https://countrysidesurvey.org.uk/

More recently Foucher et al. (2020) examined soil samples in plots for which the crop rotation history was documented since 1975; and found that last grown crop formed the dominant taxa in the eDNA inventory, alongside variable detection of past crops. The authors attributed this to different tillage practices which could affect the location of residual plant material and hence the detected eDNA. Notably this study only assessed the top 5cm of soil so other factors such as soil development processes affected by tillage practices need to be considered, which could redistribute the historic residue through the soil profile. One solution would be to sample larger volumes of soil covering a wider range of depths.

With regards to the specific sequencing methodologies, sequencing read depth coverage is also like to be a factor limiting detection, with additional cost implications. The analogy here is with a plant surveyor who is paid for throwing 50 quadrats - does the contractor want 50 locations assessed at low coverage; or focus the 50 quadrats on one location and potentially detect more rare species?

A costed run on a sequencing instrument, generates a set number of sequence reads; with the number of reads obtained per sample being dependent on how many samples are submitted within this run. More reads per sample obtained will increase the likelihood of detecting rarer members of the community - such as those sequences arising from the historic eDNA residues. Using high sequencing read depth and sequencing from larger volumes of soil would therefore potentially maximise chances of detection, but carry further cost implications.

In summary it is therefore likely that given the above considerations it should be possible to detect past crops. However pinpointing exact timescales is likely to be extremely difficult to achieve due to:

i) dependencies on soil type (chemistry, mineralogy and biology determining rates of adsorption and degradation), and

ii) the nature of the land use after crop abandonment - both with respect to tillage practice but also the post abandonment plants established, which may also affect the dilution of the historic DNA pool.

3 Summary and further considerations

Soil eDNA testing is a promising technique which on the basis of recent evidence has the potential to detect plant species grown in the last 50 years, providing consideration is paid to soil sampling volume and sequence read depth. However for current policy implementation, it is unlikely to provide an accurate estimation of the timescales with respect to accurately establishing when the previous crop was grown. In this regard, it should be considered as a method which could provide additional evidence for the presence of past crops alongside other tests/investigations.

In terms of implementation, there are barriers in that the methodologies and approach are novel, and so specialist laboratories are currently likely to be required for analyses and interpretation, although DNA analysis services are commonly available in both academic and commercial laboratories. Furthermore robust standards in both lab procedures and reporting need to be established for policy purposes, requiring initial research and development investment. Such research could for example be used to establish key relevant criteria for reporting – for example in defining quantitative indices used to indicate likelihood of past crop presence. Once validated, the costs for laboratory analyses have the potential to be low.

Estimating costs at this stage is difficult, as typically costs for molecular sequence analyses scale with sample number, and as stated above – for the prescribed purposes here there are uncertainties with respect to the amount of soil required to be sampled, as well as the amount of sequencing depth (coverage) required to detect past plant species.

The graphic of Figure-1 and details in Table-1 illustrate the techniques involved in the analyses, well as considerations of cost and practical implementation steps.

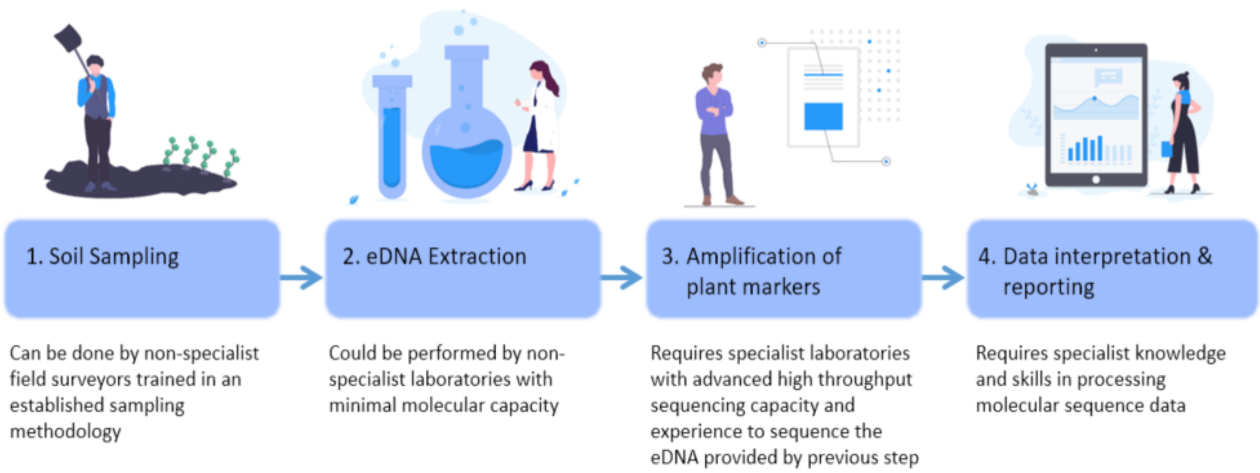


Figure 1: *Steps to implement eDNA assessment of soil.*

Table 1: *eDNA assessment of soil for detecting past land use: Implementation, development and cost considerations.*

Procedure/Step	Considerations & cost implications
1. Soil Sampling	**Soil sampling volume requirements to be standardised, based on evidence**. Larger volumes will facilitate past species detection, but will cost more with respect to sampling, processing and storage of soil material. Smaller volumes require less field effort and are easier to store – for example in a freezer.
2. eDNA Extraction	**Many protocols exist, including commercial kits, and a standard robust and cost effective methodology would need development.** For small volumes of soil (<0.5g) high throughput 96 sample kits are available allowing rapid processing of many samples (consumables costs ~ €10 per sample; ~ 2 days labour for DNA extraction of ~100 samples). Larger soil volumes will cost considerably more due to labour & need to concentrate soil extracts.
3. Amplification of plant markers	**Need to decide and develop cost effect and robust assays (marker gene) for the policy purposes.** Commercial laboratories do exist globally which could carry out the sequencing assays, or could be done with accredited research labs. Typically these services cost between €50 to €100 per sample of extracted DNA, depending on number of reads (coverage) required. The per-assay cost would likely be reduced with increased sample number.
4. Data interpretation & reporting	**Need to formulate reporting standards and metrics e.g. "Number of crop species reads detected per million sequence reads within x grams of soil" based on evidence.** Although the data processing procedures are technically involved and require programming skills, there is the potential to develop easy to use web based pipelines to process the data and report simple metrics.

To conclude, we offer the following recommendations for implementation of assays in the short-term, as well as suggestions for further research to fully evaluate the efficacy of the approach:

3.1 Recommendations for immediate application

1. Test species of interest are held in sequence databases, and have been detected with eDNA assays across a range of soils.
2. Perform *in silico* and laboratory tests to ensure current commonly used primer pairs that are used to amplify plant DNA are suitable for crop species of interest.
3. Sample larger volumes of soil to mitigate spatial localisation of past eDNA plant residues.
4. Use high read depth sequencing to ensure adequate coverage.

3.2 Recommendations for further research

5. Investigate timescales of persistence of plant eDNA across a range of soil types in laboratory assays.
6. Field calibration using sites with known past history of cultivations to quantify field efficacy.

4 Bibliography

4.1 References

Crump, S. E., Fréchette, B., Power, M., Cutler, S., Wet, G. de, Raynolds, M. K., Miller, G. H. (2021). Ancient plant DNA reveals High Arctic greening during the Last Interglacial. Proceedings of the National Academy of Sciences, 118(13). doi: 10.1073/pnas.2019069118

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology, 26(21), 5872–5895. doi: 10.1111/mec.14350

Fahner NA, Shokralla S, Baird DJ, Hajibabaei M. Large-scale monitoring of plants through environmental DNA metabarcoding of soil: Recovery, resolution, and annotation of four DNA markers. PLoS ONE. 2016;11(6).

Foucher A, Evrard O, Ficetola GF, Gielly L, Poulain J, Giguet-Covex C, et al. Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. Sci Rep. 2020 Jun 29;10(1):10502.

Levy-Booth DJ, Campbell RG, Gulden RH, Hart MM, Powell JR, Klironomos JN, et al. Cycling of extracellular DNA in the soil environment. Soil Biol Biochem. 2007 Dec 1;39(12):2977–91.

Pawlowski, J., Apothéloz-Perret-Gentil, L., & Altermatt, F. (2020). Environmental DNA: What's behind the term? Clarifying the terminology and recommendations for its future use in biomonitoring. Molecular Ecology, 29(22), 4258–4264. doi: 10.1111/mec.15643

Ruppert KM, Kline RJ, Rahman MS. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Glob Ecol Conserv. 2019 Jan;17.

Thomsen PF, Willerslev E. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Spec Issue Environ DNA Powerful New Tool Biol Conserv. 2015 Mar 1;183:4–18.

Yoccoz NG, Bråthen KA, Gielly L, Haile J, Edwards ME, Goslar T, et al. DNA from soil mirrors plant taxonomic and growth form diversity. Mol Ecol. 2012 Aug 1;21(15):3647–55.

4.2 Open access journal links

Deiner et al.: https://doi.org/10.1111/mec.14350

Fahner et al.: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0157505

Foucher et al.: https://www.nature.com/articles/s41598-020-67452-1

Pawlowski et al.: https://doi.org/10.1111/mec.15643

Thomsen et al.: https://www.sciencedirect.com/science/article/pii/S0006320714004443

BANGOR
UK Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW
United Kingdom
T: +44 (0)1248 374500
F: +44 (0)1248 362133

EDINBURGH
UK Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB
United Kingdom
T: +44 (0)131 4454343
F: +44 (0)131 4453943

LANCASTER
UK Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP
United Kingdom
T: +44 (0)1524 595800
F: +44 (0)1524 61536

WALLINGFORD (Headquarters)
UK Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB
United Kingdom
T: +44 (0)1491 838800
F: +44 (0)1491 692424

enquiries@ceh.ac.uk

www.ceh.ac.uk