BIOPLAT-EU

Promoting sustainable use of underutilised lands for bioenergy production through a web-based Platform for Europe

D2.3
Report on Tier-1 Map (incl. preliminary map)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 818083.

Project no. 818083

Project acronym: BIOPLAT-EU

Project title: Promoting sustainable use of underutilised lands for bioenergy production through a web-based Platform for Europe

Call: H2020-LC-SC3-2018-RES

Start date of project: 01.11.2018

Duration: 36 months

Deliverable title: D2.3: Report on Tier-1 map (incl. preliminary map)

Due date of deliverable: June 2020

Organisation name of lead contractor for this deliverable: *JR* (*04*)

1. table

Name	Organization
Manuela Hirschmugl, Carina Sobe	*JR*

2. table

Dissemination level		
PU	Public	
CO	Confidential, restricted under conditions set out in Model Grant Agreement	X

3. table

History			
Version	Date	Reason	Revised by
01	15/06/2020	Draft version	JR
02	30/06/2020	Final version	WIP

Table of Contents

Figures

Tables

Executive Summary

This document describes the methods and processing steps as well as the used input data for the generation of the **TIER-1 MUC map**. MUC means “marginal, underutilized and contaminated”, which was originally foreseen to be covered. However, a review of existing databases and during technical discussions within the project consortium, it turned out, that marginal lands might not be usable within the framework of BIOPLAT-EU. In BIOPLAT-EU, the frame condition was set to consider only land, which is currently not used („underutilized”) or not usable („contaminated”) for food production. This is not the case for marginal lands currently in use. Therefore, we consider **only underutilized and contaminated land** keeping in mind, that many of the marginal lands are also underutilized because of soil or economic marginality.

This deliverable D2.3 describes the more general TIER-1 approach with lower spatial resolution data, but covering whole Europe. This is in contrast to the TIER-2 approach, which is the more detailed approach applied on specific case study areas and described in Deliverable D2.4.

This deliverable covers the generation of the **contaminated land map for Europe**, for which two options were compared: first, using common thresholds for all countries and second, **using national thresholds**. The second option was used for the final map, as the results are more in line with national statistics and thus appear more plausible. Detailed accuracy assessment of the contaminated land map is not feasible within BIOPLAT-EU, as this would involve in-situ measurements throughout Europe, which was neither financially, nor time-wise planned and feasible in BIOPLAT-EU.

The mapping of **underutilized land** was based on a **remote sensing time series approach** using Landsat data with a spatial resolution of 30 m instead of the originally in the proposal foreseen MODIS data with 250 m spatial resolution. The classification was done **separately for each bio-geographical region** of Europe, as the typical appearance and signatures of underutilized land, but also for the different utilized land classes (forest, agriculture, pasture, etc.), vary considerably. Training data was generated for each region and classification was performed separately. The results were then combined.

In a next step, **areas not usable for bioenergy production** due to different reasons were **cut out**. These cut-out masks comprise: forests, water & wetlands, settlements, protected areas, steep slopes and partly agriculturally used areas. For this cutting exercise, existing **COPERNICUS layers** such as the high resolutions layers (forest, water & wetlands, settlements, agriculture) or **other open data sets** (Open Street Map, Shuttle Radar Topography Mission, Natura2000) were employed.

The **preliminary TIER-1 map is provided for the integration into the WebGIS system** together with this deliverable. The post-processing needed to make the map **fit for inclusion** into the WebGIS system is also presented. This post-processing includes smoothing and simplifying routines and minimum mapping unit application. In addition, the **required attributes** needed for the STEN modelling were generated, explained and delivered alongside the map.

The map is considered preliminary, which means, that based on feedback from project partners and users, some adjustments are still possible over the upcoming months. The preliminary results are presented through mapping examples from all over Europe, though with a focus on the case study countries (Hungary, Italy, Romania, Germany, Spain and Hungary). The examples are used for discussing some critical issues, limitations of the approach and the successfully derived map results. A proper accuracy assessment for the underutilized areas will be done after the map is finalized.

1 Introduction

The overall objective of the BIOPLAT-EU project is to *promote the market uptake of sustainable bioenergy in Europe using marginal, underutilised, and contaminated lands for non-food biomass production through the provision of a web-based platform that serves as a decision support tool*. In addition to the web-based platform, the BIOPLAT-EU project will help remove market uptake barriers of bioenergy including mainly technical, financial and legal barriers. The BIOPLAT-EU project will mobilise and involve the different stakeholders to remove these barriers by coordinating the provision of technical and financial advice by experts in these fields and by communicating with local and regional authorities in order to expedite future project's implementation.

The first objective within BIOPLAT-EU is the creation of a database of maps of MUC lands in Europe. This database is generated based on Earth Observation satellite data from Copernicus and other sources. The database is a compilation of

- results from other EU and international projects, which have already produced valuable maps, tools and information addressing sustainable bioenergy production on MUC lands;
- data compiled by the consortium from governments, public and private partners throughout the project and
- the results of the consortiums' own classification efforts in terms of time series analysis to complete the gaps.

All generated data will be included in GIS software with INSPIRE-compliant metadata files attached and transferred to a dedicated online platform (webGis system developed in WP3).

This document is a status report on the TIER-1 MUC mapping, where the preliminary map is due together with this deliverable in Month 20. This document includes all relevant descriptions on the sources, processing steps, references and the preliminary results. The so-called TIER-1 MUC map is a pan-European map of contaminated and underutilized lands. Purely marginal lands were discarded (for details and reasons please see chapter 2.1). The interactions with and dependencies from other tasks in this Work Package are illustrated in Figure 1.

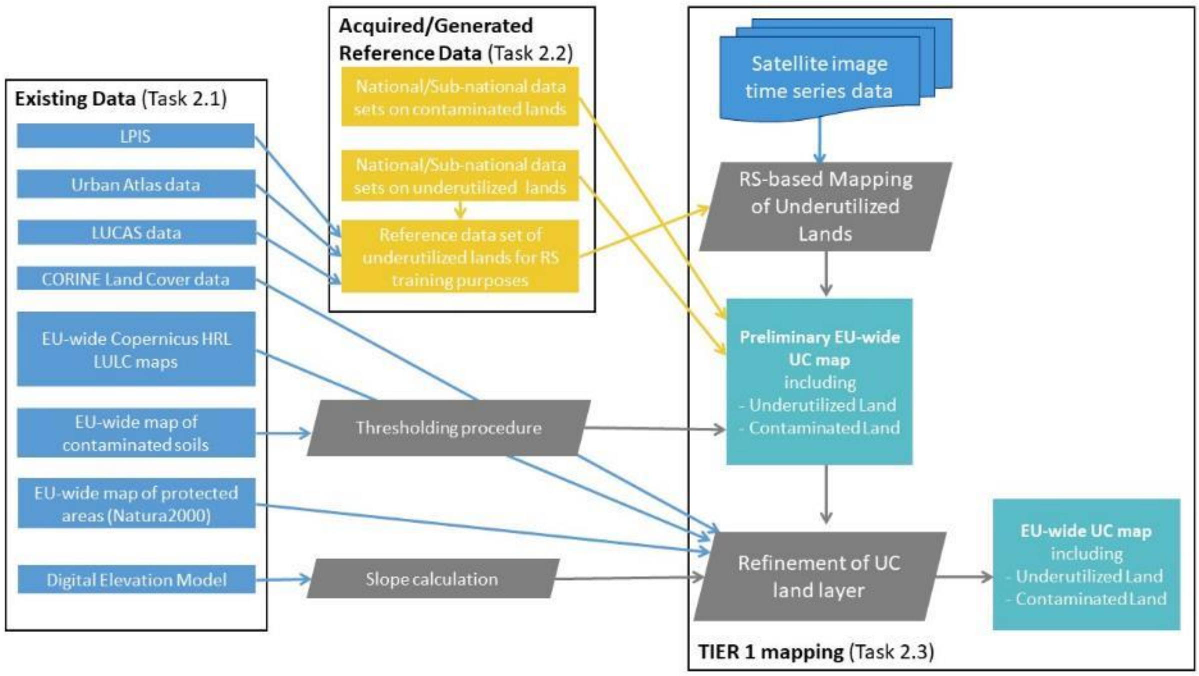


Figure 1: Work package 2 logic and interdependencies of the Tasks 2.1, 2.2 and 2.3

2 Methods

This chapter contains all methodological descriptions and the relevant data sources for the generation of the TIER-1 MUC map. It is structured along the different categories of MUC lands: marginal (Section 2.1), underutilized (Section 2.3) and contaminated (Section 2.2) land. In addition, one section is added, which explains the methods and sources for the refinement of the maps by several masks. These masks are used to cut out areas, which cannot be considered within BIOPLAT-EU for various reasons (for details see Section 2.4).

2.1 Marginal Land Mapping

Originally, also marginal lands were to be considered in the MUC land map. However, based on the review of existing databases and during technical discussions within the project consortium, it turned out, that marginal lands might not be usable within the framework of BIOPLAT-EU. In BIOPLAT-EU, the frame condition was set to consider only land, which is currently not used (compare „underutilized") or not usable (compare „contaminated") for food production. In this respect, we reviewed several marginal land databases, as e.g. provided by projects like MAGIC and SEEMLA or from the literature (Sallustio et al., 2018). This review showed, that large parts of the mapped lands, which are (correctly) considered marginal due to economic or soil parameters, are currently used for food production. This is extremely evident in southern Italy, where olive groves and vineyards stretch over large areas of marginal lands, see Figure 2 (source: SEEMLA). Having to consider this, we decided to retain only contaminated and underutilized lands in all follow-up mapping (both on TIER-1 and TIER-2). However, this does not mean, that no marginal lands are included in the MUC land map. All marginal lands, which have not been used during the past five years because of e.g. insufficient economic benefit, would still be included within the category „underutilized land". Such an example is shown in Marginal lands, which are contaminated in some way, are included under the category „contaminated lands".

Figure 2: Marginal lands used for food production (mainly olives and wine) in Italy

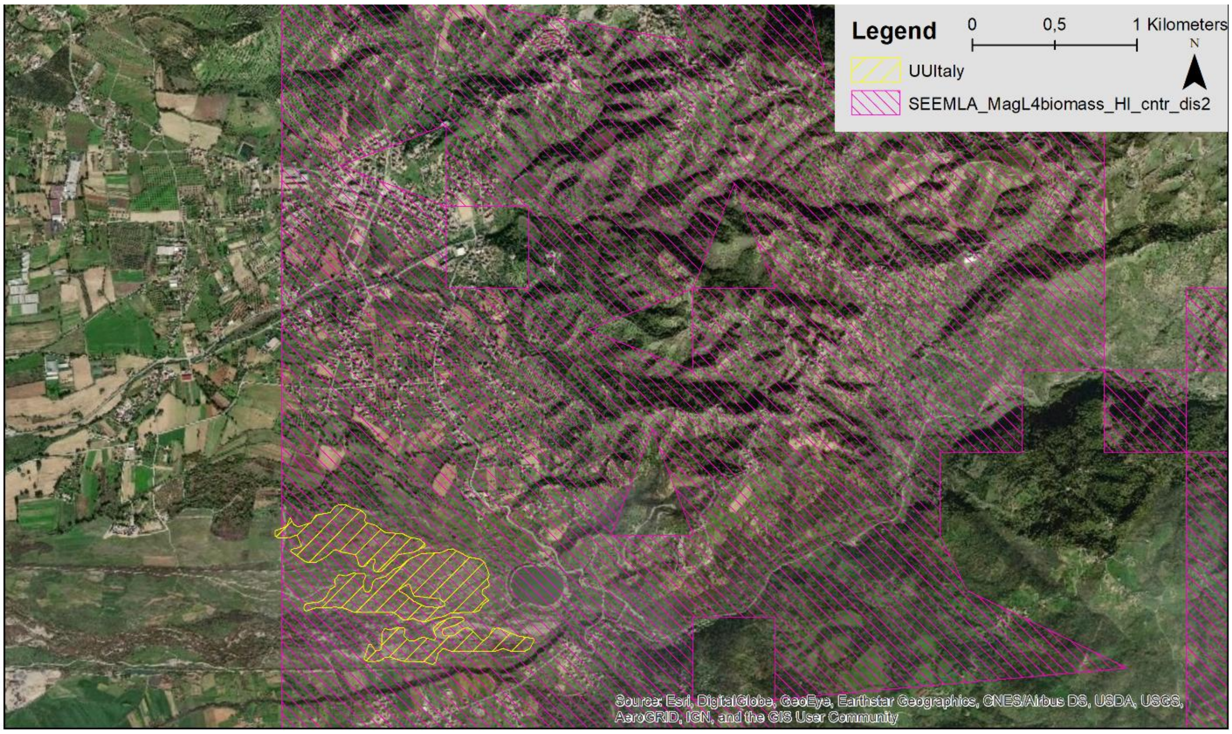


Figure 3: Marginal lands from SEEMLA compared to underutilized land in BIOPLAT-EU

2.2 Contaminated Land Mapping

For the generation of an EU-wide contaminated land layer, we pursued two options: The first attempt was to collect national maps to compile a full coverage of Europe (bottom-up approach). However, due to limitations in availability of these national maps, this option turned out not to be feasible. Although most member states report statistics on contaminated lands (shares of total land), many countries either do not have or do not share the underlying spatial data sets due to legal restrictions. In many cases (e.g. Hungary), there is only point-wise data available. In other countries, such as Romania, the official contaminated land layer is still under evaluation and not finally and officially released. These limitations led us to the second option: a top-down approach using an EU-wide map of contaminations, which we derived by the Joint Research Centre (JRC) in the "Heavy metals in soils" product based on LUCAS 2009 heavy metal (HM) data. It is clear, that this data set is not as accurate and as detailed as national maps, however in order to fulfil the requirements to provide an EU-wide map of contaminated lands, this was the only feasible option. Unfortunately, Ukraine is not covered with this data set. It should be mentioned, that for the TIER-2 maps of the case study areas, available national and regional data sets are replacing the TIER-1 results. In addition, for countries with available national maps, such as Italy, we included both layers in the TIER-1 map.

The JRC map of HM in soils is available at https://esdac.jrc.ec.europa.eu/content/maps-heavy-metals-soils-eu-based-lucas-2009-hm-data-0#tabs-0-description=1. It has a spatial resolution of 1x1 km and covers 27 EU member states (not including Croatia). Maps of nine different heavy metals are provided: Arsenic, Cadmium, Chromium, Cobalt, Copper, Mercury, Nickel, Lead, Manganese and Antimony. For each of the heavy metals, thresholds had to be defined to separate contaminated from non-contaminated soils. The threshold values represent the amount of heavy metals in soils, above which the use of the soil for food and fodder are not allowed/advisable. The relevant EU directive (Council of the European Union, 2002) gives only ranges of values (see Table 1 "EU directive") rather than a specific threshold value. Previous studies (Toth et al., 2016) used Finnish thresholds for whole Europe, as these thresholds are well in line with EU-directive (see Table 1, "JRC map" and "EU directive"). The possibilities to derive a European map of contaminated lands are shown in Figure 6. Finally, the top-down approach with option B was used.

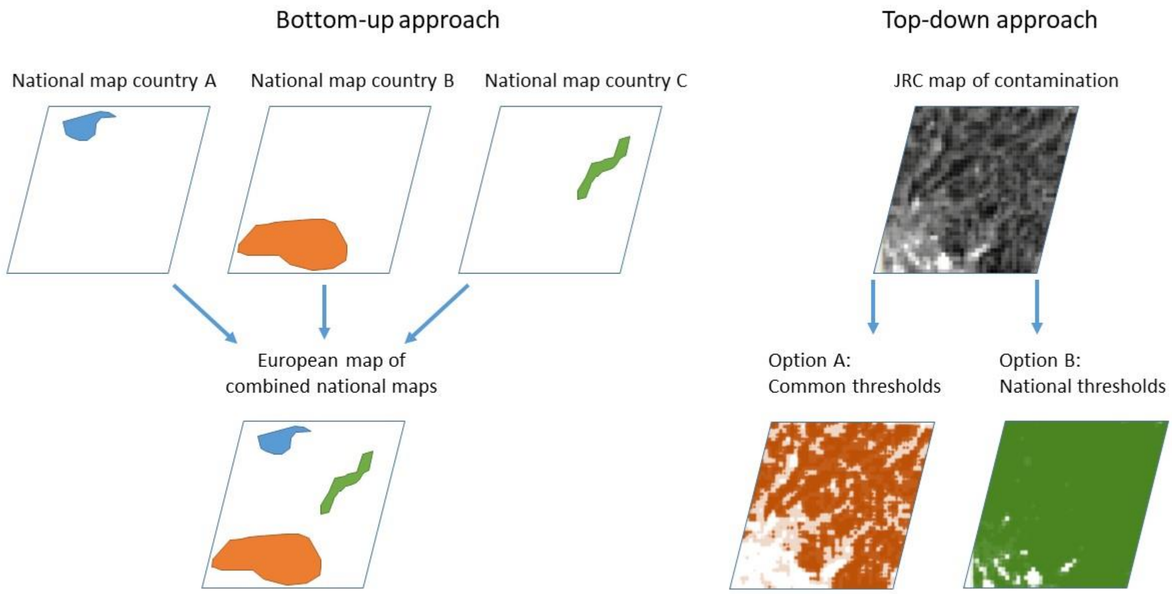


Figure 4: Different possibilities to derive a European map of contaminated lands

At this point, it is worth mentioning, that national legislations often have divergent regulations. In Germany for example, the German Soil Protection Act (Bundesministerium für Justiz und Verbraucherschutz, 1998) certainly provides the possibility of defining protective and restrictive measures (§8 BBodSchG). These enable, for example, an adaptation of the utilization and management of the affected soils (§5 BBodSchVO, Bundesministerium für Justiz und Verbraucherschutz 1999), in particular through restrictions on the use, to prevent further accumulation of pollutants (LUBW, 2018). Measured values, that have been defined for individual hazardous substances must be taken into account. The type of usage restrictions depends on the specific situation on site, in particular on the type of current use and the type of load.
Furthermore, the Feed Regulation (Bundesministerium für Justiz und Verbraucherschutz, 2010) plays a role, in which the use of so-called 'undesirable substances' in animal nutrition is excluded. It refers to Directive 2002/32/EC of the European Parliament and the Council of May 7, 2002 on undesirable substances in animal nutrition (Council of the European Union, 2002). Annex I of the directive lists various substances and their maximum permitted levels in products intended for animal nutrition. The European Regulation (EC) No. 1881/2006 laying down the maximum levels for certain contaminants in food regulates the use of plants or parts of plants for human consumption (European Commission, 2006). In Germany, the regulation on the limitation of contaminants in food (Contaminants Regulation - KmV) also applies (Bundesministerium für Justiz und Verbraucherschutz, 2010). Consequently, it may be possible to grow biomass on contaminated land, which, however, may not be used in animal nutrition or for human consumption. Regarding the substantial use of biomass (e.g. biopolymers), there may also be regulations on the permitted maximum levels of certain substances in parts of plants that may only have been agreed between the buyer and the provider of the biomass, for example to guarantee certain quality standards.
Worldwide, more than 400 plant species are identified, which take up and translocate metal contaminants (Zn, Ni, Cd and Pb) via the roots into the aboveground biomass far beyond the physiological optimum and at a level 100-150-fold greater than common plants without yield reduction (Brooks 1979, Chaney et al 2007, Paz-Ferreiro et al. 2014).

This means, that the absorption capacity of individual plant species play a crucial role: the same area with a certain soil contamination could safely be used to cultivate a specific crop for food production, but a different crop would exceed the allowed thresholds. Examples of plant species with a high to very high accumulation rate of heavy metals in aboveground biomass are sunflower and amaranth. Moreover, also water balance, climate and other parameters influence the absorption capacity. Such a detailed modelling per plant species and for whole Europe with all its climatic variations is clearly far beyond the scope of BIOPLAT-EU. Therefore, we limit our approach to soil contaminations aware of possible shortcomings entailed in this approach.

We performed the analysis with the Finnish thresholds. The results revealed very large amounts on contaminated lands in some countries, which are not in line with national legislation and the project partners' understanding of the situation in their countries. Therefore, we decided to compare these results to the results when using national thresholds instead. For this comparison, we collected national thresholds from official sources. Table 1 provides the thresholds for all heavy metals and all countries. Table 2 lists the countries and relevant sources for these thresholds.

Table 1: Thresholds for heavy metal concentrations in soils in different countries

	Andorra	Austria	Belgium	Bulgaria	Croatia	Cyprus	Czech Rep.	Denmark	Estonia	Finland	France	Germany
Heavy metal type	AND	AUT	BEL	BGR	HRV	CYP	CZE	DNK	EST	FIN	FRAN	DEU
	threshold (mg/kg)											
Arsenic (As)	5	20	27	5	5	5	30	10	5	5	19	50
Cadmium (Cd)	1	1	1.2	1	1	1	1	0.5	1	1	2	1.5
Chromium (Cr)	100	100	100	100	100	100	200	30	100	100	150	100
Copper (Cu)	100	100	100	100	100	100	100	40	100	100	100	60
Mercury (Hg)	0.5	1	1	0.5	0.5	0.5	0.8	0.5	0.5	0.5	1	1
Nickel (Ni)	50	50	50	50	50	50	80	15	50	50	50	50
Lead (Pb)	60	100	100	60	60	60	140	40	60	60	100	100
Zinc (Zn)	200	150	200	200	200	200	200	100	200	200	300	*200*
Cobalt (Co)	20	50	20	20	20	20	50	20	20	20	120	20
Manganese (Mn)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Antimony (Sb)	2	2	2	2	2	2	2	2	2	2	50	2
Vanadium (V)	100	50	100	100	100	100	220	100	100	100	280	100
Molybdenum (Mo)	n/a	10	n/a	n/a	n/a	n/a	5	2	n/a	na	100	n/a

	Greece	Hungary	Ireland	Italy	Latvia	Liechten-stein	Lithuania	Luxem-bourg	Nether-lands	Norway	Poland	Portugal	Romania
Heavy metal type	GRC	HUN	IRL	ITA	LVA	LIE	LTU	LUX	NLD	NOR	POL	PRT	ROU
	threshold (mg/kg)												
Arsenic (As)	5	15	5	20	5	5	10	5	29	5	20	5	25
Cadmium (Cd)	3	1	1	1.5	1	1	3	3	0.8	1	1	4	5
Chromium (Cr)	100	75	100	150	100	100	100	200	100	100	150	300	300
Copper (Cu)	140	75	50	100	100	100	100	140	36	50	150	200	200
Mercury (Hg)	1.5	0.5	1	1	0.5	0.5	1.5	1.5	0.3	1	2	2	2
Nickel (Ni)	75	40	30	75	50	50	75	75	35	30	100	110	150

Lead (Pb)	300	100	50	100	60	60	100	300	85	50	100	450	100
Zinc (Zn)	*300*	200	150	300	200	200	300	300	140	150	300	450	600
Cobalt (Co)	20	30	20	20	20	20	30	20	9	20	20	20	50
Manganese (Mn)	n/a	n/a	n/a	n/a	n/a	n/a	1500	n/a	n/a	n/a	n/a	n/a	n/a
Antimony (Sb)	2	2	2	10	2	2	10	2	3	2	2	2	20
Vanadium (V)	100	100	100	90	100	100	150	100	42	100	100	100	200
Molybdenum (Mo)	n/a	7	n/a	n/a	n/a	n/a	5	n/a	3	n/a	10	n/a	n/a

	Slovakia	**Slovenia**	**Spain**	**Sweden**	**Switzerland**	**Ukraine**	**United Kingdom**	**Vatican City**	**Europe (JRC map)**	**EU directive**
Heavy metal type	**SVK**	**SVN**	**ESP**	**SWE**	**CHE**	**UKR**	**GBR**	**VAT**		
	threshold (mg/kg)									
Arsenic (As)	30	5	5	15	5	n/a	50	5	5	n/a
Cadmium (Cd)	5	1	3	0.4	0.8	3 (5)*	3	1	1	1 - 3
Chromium (Cr)	250	100	150	60	75	100 (300)	400	100	100	n/a
Copper (Cu)	100	100	210	40	50	100 (200)	80	100	100	50 - 140
Mercury (Hg)	2	0.5	1.5	0.3	0.8	n/a	1	0.5	0.5	1 - 1.5
Nickel (Ni)	100	50	112	30	50	50 (70)	50	50	50	30 - 75
Lead (Pb)	150	60	300	40	50	100 (150)	300	60	60	50 - 300
Zinc (Zn)	500	200	450	100	200	300 (500)	300	200	200	150 - 300
Cobalt (Co)	20	20	20	30	20	30 (50)	20	20	20	n/a
Manganese (Mn)	n/a	n/a	n/a	n/a	n/a	1500(3000)	n/a	n/a	n/a	n/a
Antimony (Sb)	2	2	2	2	2	n/a	2	2	2	n/a
Vanadium (V)	200	100	100	120	100	n/a	100	100	100	n/a
Molybdenum (Mo)	40	n/a	n/a	n/a	n/a	4 (5)	n/a	n/a	n/a	n/a

* Value outside parentheses is for Forest-Steppe zone. Value in parentheses is for Steppe zone.

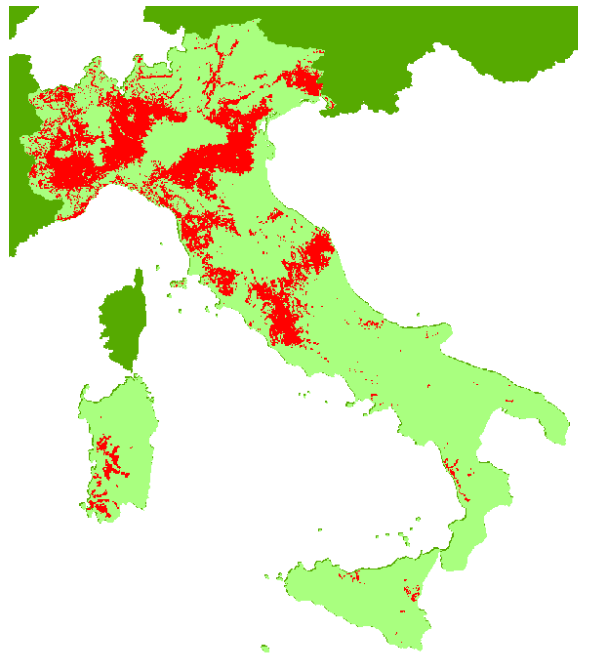
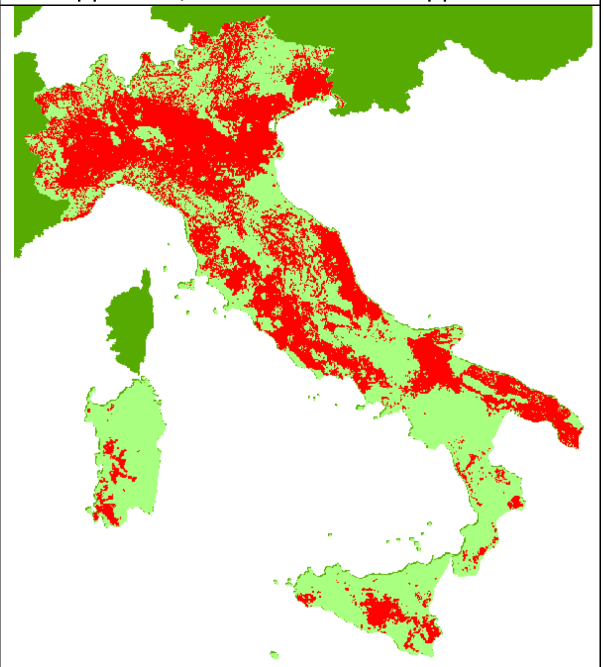
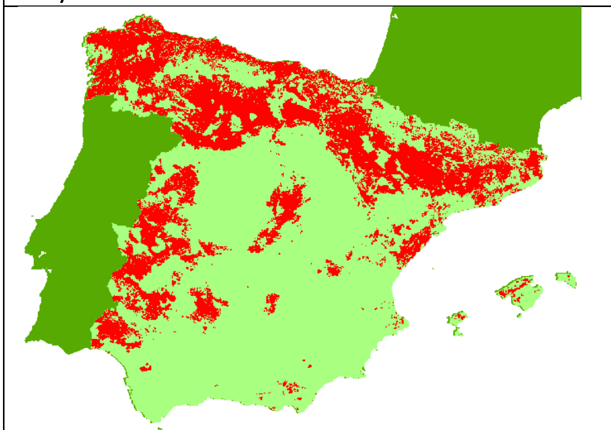
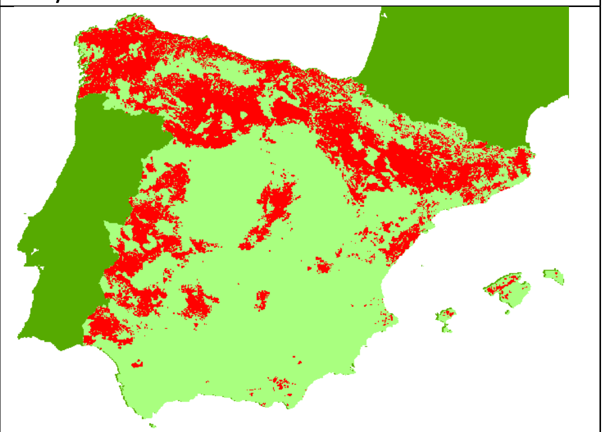
Table 2: Sources for heavy metal concentration thresholds in soils in different countries

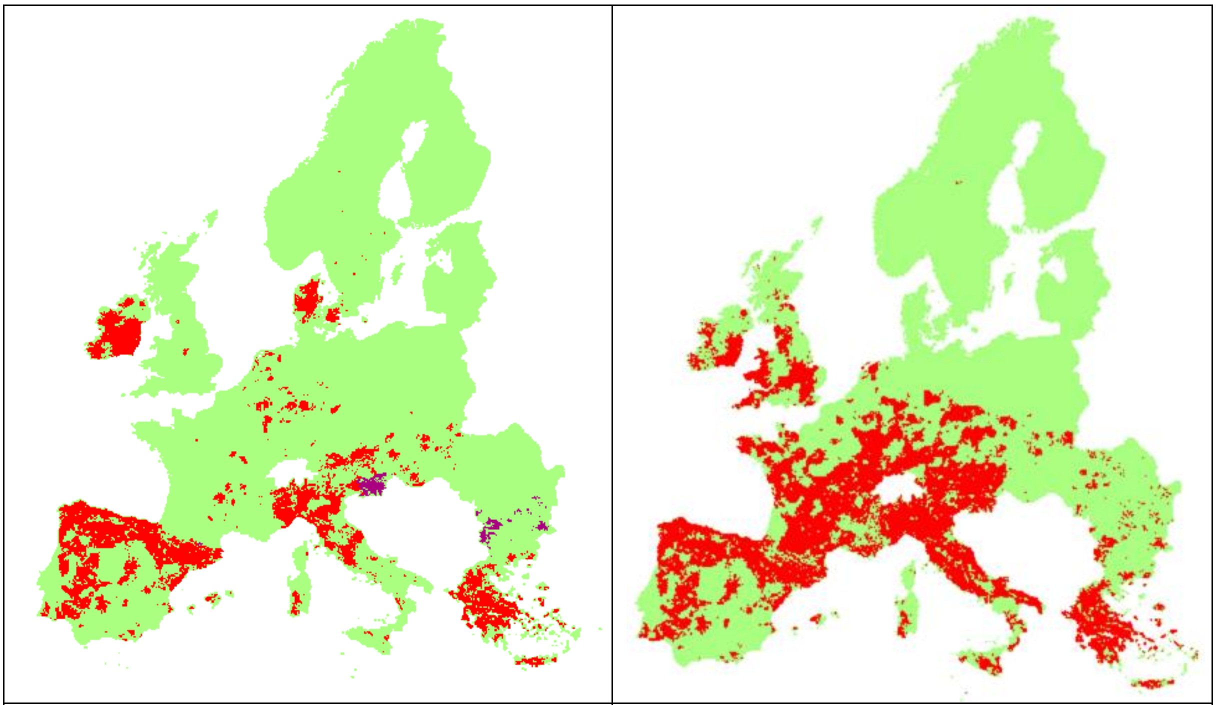
<table>
<tr><th>Country</th><th colspan="13">Heavy Metal</th></tr>
<tr><th></th><th>Arsenic</th><th>Cadmium</th><th>Chromium</th><th>Copper</th><th>Mercury</th><th>Nickel</th><th>Lead</th><th>Zinc</th><th>Cobalt</th><th>Manganese</th><th>Antimony</th><th>Vanadium</th><th>Molybdenum</th></tr>
<tr><th></th><th>(As)</th><th>(Cd)</th><th>(Cr)</th><th>(Cu)</th><th>(Hg)</th><th>(Ni)</th><th>(Pb)</th><th>(Zn)</th><th>(Co)</th><th>(Mn)</th><th>(Sb)</th><th>(V)</th><th>(Mo)</th></tr>
<tr><td>Austria</td><td colspan="9">Amlinger et al 2004</td><td>n/a</td><td>Carlon 2007</td><td>UBA 2001</td><td>Amlinger et al 2004</td></tr>
<tr><td>Belgium*1</td><td colspan="8">Amlinger et al 2004</td><td colspan="5">n/a</td></tr>
<tr><td>Czech Rep.</td><td>Carlon 2007</td><td colspan="9">Vacha & Sanka 2014</td><td>Carlon 2007</td><td>Vacha & Sanka 2014</td><td>Carlon 2007</td></tr>
<tr><td>Denmark</td><td>Carlon 2007</td><td colspan="7">Amlinger et al 2004</td><td colspan="5">n/a</td></tr>
<tr><td>Finland</td><td colspan="9">Carlon 2007</td><td>n/a</td><td colspan="2">Carlon 2007</td><td>n/a</td></tr>
<tr><td>France</td><td>Carlon 2007</td><td colspan="7">Amlinger et al 2004</td><td>Carlon 2007</td><td>n/a</td><td colspan="2">Carlon 2007</td><td>Carlon 2007</td></tr>
<tr><td>Germany*[1]</td><td>Carlon 2007</td><td colspan="7">Amlinger et al 2004</td><td colspan="5">n/a</td></tr>
<tr><td>Greece</td><td>n/a</td><td>Amlinger et al 2004</td><td>n/a</td><td colspan="5">Amlinger et al 2004</td><td colspan="5">n/a</td></tr>
<tr><td>Hungary</td><td colspan="9">Sipos & Poka</td><td colspan="3">n/a</td><td>Sipos & Poka</td></tr>
<tr><td>Ireland</td><td>n/a</td><td>Amlinger et al 2004</td><td>n/a</td><td colspan="5">Amlinger et al 2004</td><td>n/a</td><td colspan="4">n/a</td></tr>
<tr><td>Italy</td><td>Carlon 2007</td><td>Amlinger et al 2004</td><td>Carlon 2007</td><td colspan="5">Amlinger et al 2004</td><td>Carlon 2007</td><td>n/a</td><td colspan="2">Carlon 2007</td><td>n/a</td></tr>
<tr><td>Lithuania</td><td colspan="13">Carlon 2007</td></tr>
<tr><td>Luxembourg</td><td>n/a</td><td colspan="7">Amlinger et al 2004</td><td colspan="5">n/a</td></tr>
<tr><td>Netherlands</td><td colspan="8">Carlon 2007</td><td>Dutch Standards</td><td>n/a</td><td>Dutch Standards</td><td>Carlon 2007</td><td>Dutch Standards</td></tr>
<tr><td>Norway</td><td>n/a</td><td colspan="7">Witter 2009</td><td colspan="5">n/a</td></tr>
</table>

<table>
<tr><td>Poland</td><td>Carlon 2007</td><td>n/a</td><td>Carlon 2007</td><td colspan="4">n/a</td><td>Carlon 2007</td></tr>
<tr><td>Portugal*[2]</td><td>n/a</td><td colspan="2">Amlinger et al 2004</td><td>n/a</td><td colspan="4">n/a</td></tr>
<tr><td>Romania</td><td colspan="5">Moldoveanu 2014</td><td colspan="2">Romanian Ministry of Environment, order No.756/1997</td><td>n/a</td></tr>
<tr><td>Slovakia</td><td colspan="4">Carlon 2007</td><td colspan="2">n/a</td><td colspan="2">Carlon 2007</td></tr>
<tr><td>Spain</td><td>n/a</td><td colspan="2">Amlinger et al 2004</td><td colspan="5">n/a</td></tr>
<tr><td>Sweden</td><td>Carlon 2007</td><td colspan="2">Amlinger et al 2004</td><td>Carlon 2007</td><td colspan="2">n/a</td><td>Carlon 2007</td><td>n/a</td></tr>
<tr><td>United Kingdom</td><td>Amlinger et al 2004</td><td colspan="2">Witter 2009</td><td colspan="5">n/a</td></tr>
</table>

*[1] In case of two or more values, the value closer to the EU limit is applied.
*[2] Values depending on pH value of soil. Higher values are applied.

The comparison revealed varying differences in the individual countries. Spanish thresholds on the one hand are very similar to the Finnish ones, thus also the resulting maps show comparable extents. On the other hand, the maps for Germany and Italy reveal very large differences. The JRC approach (using Finnish thresholds) yields much larger shares of contaminated land than using national thresholds in these cases. In contrast, for Denmark, the national thresholds are stricter than the Finnish ones. Here, using the Finnish thresholds led to smaller areas. The following figure (Figure 5) shows these effects. For each of these maps, first, the contaminated land with the respective thresholds was calculated. In a second step, the same cut-out mask was applied. Areas of steep slopes, protected areas, areas covered by forests, water and wetlands or settlements were removed using these cut-out masks. All details on the content and generation of the cut-out masks are given in Chapter 2.4.

National thresholds applied	JRC approach / Finnish threshold applied
	
Italy national threshold: ~ 27 400 km^2	Italy JRC threshold: ~ 60 800 km^2
	
Spain national threshold: ~ 82 200 km^2	Spain JRC threshold: ~ 82 200 km^2



Europe national thresholds: ~ 198 600 km^2	Europe JRC threshold: ~ 458 400 km^2
Contaminated land with EU threshold as no national thresholds available	No contamination or cut out due to criteria above Contaminated land

Figure 5: Comparison of JRC (Finish) threshold and national thresholds for selected countries and Europe

Due to the large difference in extent (more than double) and the clear view of the local partners and counterparts, that the national thresholds are to be preferred, we decided to go forward using the national thresholds. Only for countries, where no official information on national thresholds could be found, the Finnish thresholds were used (e.g. Slovenia, Bulgaria).

2.3 Underutilized Land Mapping

For the detection of land, which is currently not in use („underutilized"), we adopted the official FAO time span of abandoned farmland. In FAO's World Census of Agriculture, FAO 2020, Art. 8.2.24 states: *„Land remaining fallow for too long may acquire characteristics requiring it to be reclassified, such as "permanent meadows and pastures" (if used for grazing), "forest and other wooded land" (if overgrown with trees), or "other land" (if it becomes wasteland). A maximum idle period should be specified –* ***five years is usually suitable****. Land cultivated on a two- or three-year rotating basis is considered to be fallow if it was not cultivated during the reference year. Land temporarily fallow should be distinguished from land abandoned by shifting cultivation; the former is part of the holding, whereas the latter is not."* Since we do not only consider abandoned farmland, but all underutilized land, we extended the definition to: land, which does not show any signs of human use for the past five years. In order to assess the existence of signs of human use, we employ time series of remotely sensed data from Copernicus and other Earth Observation programs.

2.3.1 Remote Sensing Background

In the frame of Earth Observation, we distinguish different definitions of resolution. These are explained with examples below.

- **Spatial resolution** is the area on the ground, which is represented by one pixel in the image. High spatial resolution means a small size on the ground is covered by one pixel. Sentinel-2 data has a high spatial resolution of 10 m (in the visible bands). This means, a 10 by 10 m patch on the ground is represented by one pixel in the image and has one spectral value per band. MODIS data on the other hand has a low spatial resolution of 250 m. Figure 6 compares images of the same area as mapped by these two sensors (MODIS vs. Sentinel-2).

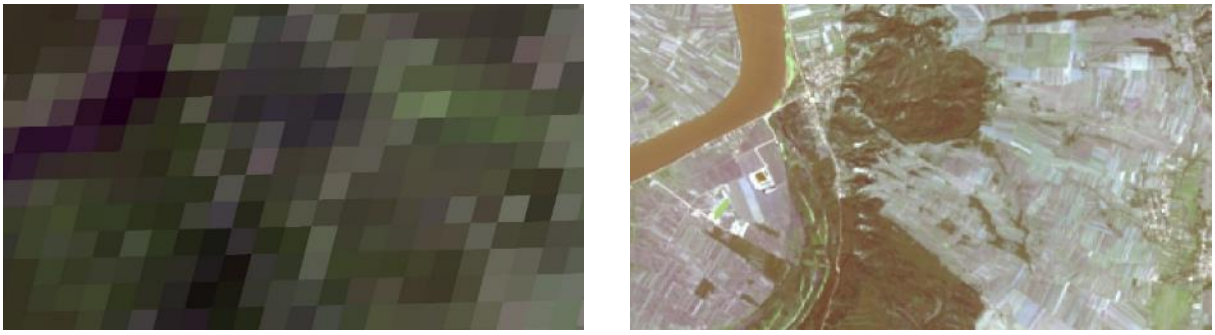


Figure 6: Comparison of low spatial resolution MODIS data (left) and high spatial resolution Sentinel-2 data (right)

- **Spectral resolution** is the ability to resolve features in the electromagnetic spectrum. Higher spectral resolution means more separate parts of the spectrum (i.e. bands) are acquired. Some sensors cover only the visible light range; others cover the spectrum far into the infrared range. The near and shortwave infrared part of the spectrum is sensible to photosynthetic activity and thus important for vegetation mapping.

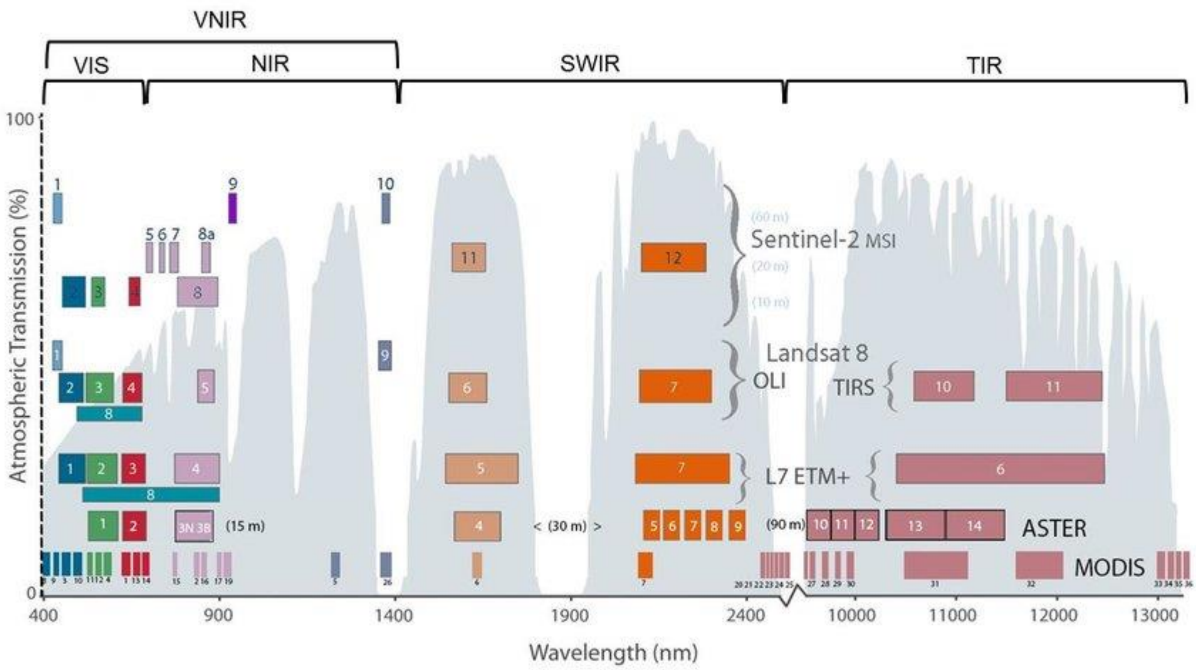


Figure 7: Spectral resolution of currently available optical satellite sensors (from Friedl, 2019)

- **Temporal resolution** is the repeat rate of a satellite sensor system. It can also be explained as the time between observations of the same spot on the ground. High temporal resolution means a short time between observations, e.g. from daily to bi-weekly. The term „time series analysis" usually requires such high temporal resolution data as input.

These three resolutions show some logical interconnections. When the spectral resolution is high (many bands), each band covers only a small part of the electromagnetic spectrum. In order to receive sufficient energy from that narrow band, the size of the pixel has to be large (lower spatial resolution). If the spatial resolution is low and the corresponding covered size per pixel is high, the area covered on the ground is large. When the satellite overpasses the next time and covers a large area again, there is a lot of overlap between the two images. This overlap increases the temporal resolution. Thus, typically, low spatial resolution data has higher temporal and high spectral resolution (MODIS) compared to high spatial resolution data (Landsat, Sentinel-2).
Sentinel-2 is a specific case, as it is a constellation of two synchronously working satellites (Sentinel-2A and Sentinel-2B) with the same sensor. Therefore, it can provide both high spatial resolution and quite high temporal resolution (5 days repeat rate).

Previous works in classification of underutilized or abandoned land using Earth Observation data were mainly done in Eastern Europe after 2000, where large agricultural areas had been abandoned. Studies used either high-temporal & low spatial resolution data such as MODIS (Alcantara et al., 2013; Estel et al., 2015; Estel et al., 2016; Löw et al., 2018) or low temporal & high spatial resolution data for selected countries (e.g. Slovakia: Szatmari et al., 2018; Czech Republic: Soukup et al., 2009). Initially, as detailed in the proposal, the TIER-1 mapping of underutilized areas in Europe was also foreseen to be based on MODIS data. This satellite system with a spatial resolution of 250 - 500 m is well suited to identify large and homogeneous area, which are not in use. However, reviewing the small structured agricultural lands in large parts of Europe, this approach turned out to be very coarse. Many potentially interesting areas would not be detectable. Although processing higher resolution data was very time consuming and much more work than anticipated, we decided to proceed with a more detailed approach. In order to fulfil the definition of underutilized land, we employed Landsat 8 data with a spatial resolution of 30 m, as this kind of data was already available for the past five years by the beginning of the project work. The time series was processed using Google Earth Engine.

2.3.2 Training Data

Training data is a crucial component in any remote sensing-based classification. It is important in order to „train” the classifier the typical characteristics of underutilized land as well as those of other land cover types to distinguish it. Representative training samples are key to a successful classification of the relevant topic. In case of underutilized land, we have to consider that these lands can have a wide range of different characteristics depending on climate, soil and topography. Therefore, being “representative” is even more crucial. Figure 8 highlights the wide range of characteristics of areas considered as “underutilized” in different parts of Europe. The land cover can be bare soil, grass-covered or contain shrubs and even trees. All these land cover types are different and have to be present in the reference data in order to be correctly classified.

In addition to training data for the class „underutilized”, also training data for the various „utilized” classes is needed. These training data sets were generated with the help of existing data sets such as LUCAS (point data), Copernicus HRLs for forest, settlement and water & wetlands and CORINE land cover data.

Figure 8: Different characteristics of underutilized lands in different parts of Europe

From the partners' inputs through the platform (see Deliverable D2.2), we received inputs for several countries like Ukraine, Germany, Lithuania, etc. However, this was not sufficient data for the training, as several geographic areas were completely missing. Therefore, we also explored alternative sources for collecting data to improve the amount and variety of training data.

Existing EU-wide data sets can partly be used for training. These are:

- LUCAS points, where land use is 410, 420 or 112 in the years 2012, 2015 and 2018 and land cover is not water, wetland or forest
- Urban Atlas polygons of classes 32000, 33000 and 13400
- LPIS data depending on availability

Sufficient training data has to be available for each bio-geographical regions of Europe. We processed the EU plus Norway, Switzerland, the Balkan region and Ukraine, but excluded the Atlantic islands like Canary Islands, Madeira and the Azores, Iceland and Spitsbergen, as well as Malta and Cyprus in the Mediterranean Sea. The European Environment Agency (EEA) has defined these regions and is using them in all official documents. The type and extent of the bio-geographical regions of Europe is depicted in Figure 9.

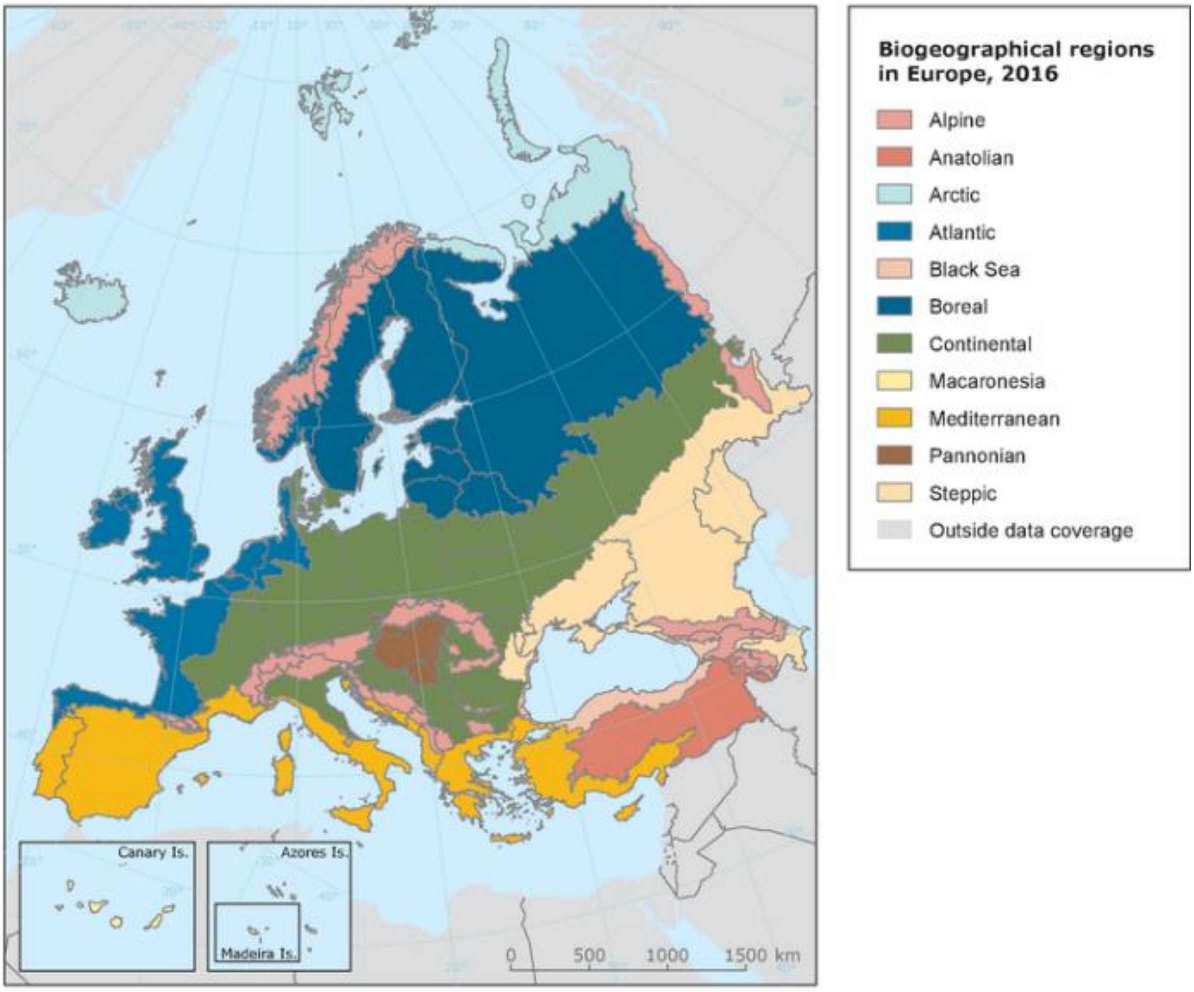


Figure 9: Biogeographical regions of Europe (Source: EEA)

2.3.2.1 LUCAS points for training

The LUCAS points have to be checked visually with Google Earth data and transformed into polygons in order to be used for training within BIOPLAT-EU. One example of a point to be rejected is shown in Figure 10: a formerly underutilized area has been built up and can therefore not be used for training. Another example is shown in Figure 11: this plot is usable for training, as it is an abandoned land in the transition to forest, but not qualified as forest land yet. In order to use the area this point is representing, we drew a polygon to limit the extent of the area with the properties given in the point database. The white line shown in Figure 11 (bottom) illustrates this area. This way, we generated 1236 polygons throughout Europe as training data.

2012

2014

2017

Figure 10: Appearance of a LUCAS point over time – not usable for training

2014

2017

2018

Figure 11: Appearance of a LUCAS point over time – usable for training

2.3.2.2 *Urban Atlas data for training*

Similar to LUCAS points, also data from Urban Atlas can be used to generate reference areas. The advantage here is the already existing polygon structure. An example of UA data is shown in Figure 12. It is an area labelled as "13400 Land without current use". Checking the time series in Google Earth, it is clearly underutilized land in the sense of BIOPLAT-EU and can therefore be used for training. However, since the Urban Atlas is focussing on function/al urban areas, many parts are within or very close to settlements. These areas are often set aside for future developments (industrial, infrastructure or residential buildings) and therefore most unlikely to be used for energy production. Thus, we used only few examples for training from Urban Atlas.

Figure 12: UA land without current use in 2012 and check with time series

2.3.2.3 *LPIS data for training*

LPIS data can be a useful data source (see Deliverable D2.1). Multi-temporal LPIS data, as for example available for Czech Republic, can be used to select areas attributed to be "fallow land" consecutively. They also have to be visually checked for compliance in the satellite data time series. An example is shown in Figure 13 for an area, which is attributed to have been fallow between 2015 (first available data) and 2018 (last available data). In many other countries, such as Austria, data is only available for one year. Although this data sets contains areas termed as "fallow", it can be fallow land only during this year in the frame of normal shifting cultivation practise. Thus, such data is not suitable to generate training data for underutilized areas with our definition of land being out of use for the past five years.

	20.05.2014
	01.08.2017
	13.08.2018

Figure 13: LIPS data of Czech Republic attributed as „fallow land" between 2015 and 2018 checked with time series

2.3.3 Classification Method

The classification is conducted using the Google Earth Engine (GEE; Gorelick et al., 2017). This is an online, cloud-processing engine, which can be used free of charge for research projects. The big advantages of GEE are: its simplicity; the availability of large amounts of ready to use input data and the availability of pre-defined functions from a wide user group. At the start of BIOPLAT-EU project, Sentinel-2 data was not available in GEE. Landsat data, however, was available for the past five years and already pre-processed to surface reflectance (SR) data. Thus, we decided to focus the main processing of TIER-1 on Landsat data.
The biggest disadvantage of optical sensors is the fact that they cannot "see" through clouds. Hence, optical images might have cloudy pixels that do not represent the reflectance from the earth's surface. When performing image time series analysis, this leads to outliers in actual trajectory and, subsequently, misclassifications might occur. Hence, an additional pre-processing step to mask out cloudy pixels was applied. All Landsat 8 SR data have an additional quality band (pixel_qa) that contains information about clouds, cloud and cirrus cloud confidence, cloud shadow, snow/ice and water (Foga et al., 2017). This band is integrated in a pre-defined GGE cloud masking algorithm that we used within our processing chain to eliminate clouds in images.
Initially, several individual Landsat bands and indicators were calculated. After preliminary quality analyses, the decision was made to use the following ones within the final algorithm:

- B3 (green)
- B4 (red)
- B5 (nir)
- B6 (swir)
- NDVI (Normalized Difference Vegetation Index)
- NDII5 (Normalized Difference Infrared Index)
- MCARI (Modified Chlorophyll Absorption in Reflectance Index)
- MSAVI (Modified soil-adjusted vegetation Index)

From each of these bands, monthly images were calculated to reduce the amount of data:

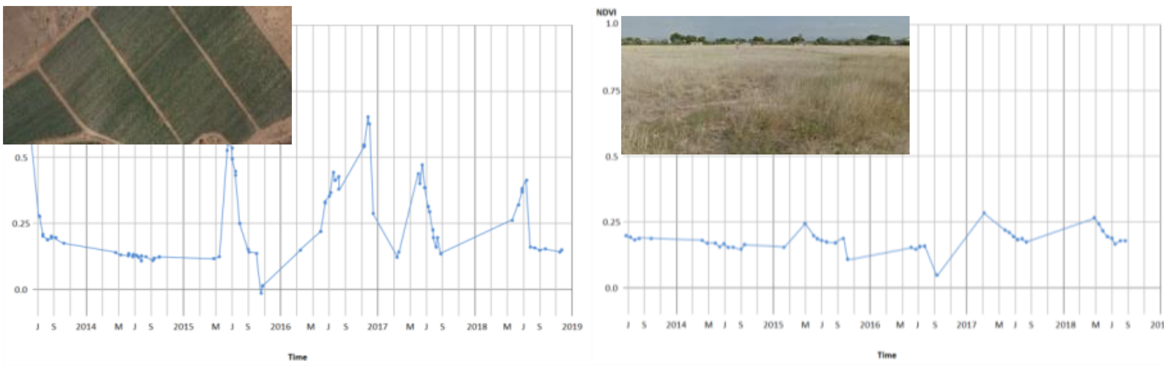
- Minimum pixel value: B3, B5, NDVI, NDII5, MSAVI, MCARI
- Maximum pixel value: B4, B6

The following temporal statistics from May – September for the 5 year period were calculated for each band and index:

- Minimum
- Maximum
- Standard Deviation
- Percentile 10 and 90

This resulted in a data set of 40 input features for the random forest (RF) classification. RF is a classification method, which belongs along with other boosting and bagging methods as well as classification trees in general to the ensemble learning methods, which generate many classifiers and aggregate their results to calculate their response (Liaw and Wiener, 2002; Horning et al., 2010; Li et al., 2016). The random forest algorithm learns the relationship between predictor and response data and can handle continuous, categorical and binary data sets (Ali et al., 2012; Horning et al., 2010; Grinand et al., 2013). The random forest algorithm offers a good prediction performance and is computationally effective but sensitive to the sample design. Colditz (2015) tested the impact on several sampling designs on decision tree algorithms and recommends the area-proportional allocation to achieve the best classification results because classes occupying larger areas need more training samples. Other authors (Mellor et al., 2015; Ali et al., 2012) found out that the random forest algorithm is less sensitive to outlier training samples or noisy data. Furthermore, Dalponte et al. (2013) proposed that the algorithm fails to cope with imbalanced training data tending to favour the most representative class at the expense of the minority class. Thus, at each sample selection at each node during the tree construction fewer samples of the minority class are chosen. Also, the size of the trainings data has an impact on the classification accuracy. Colditz (2015) recommend a sample size of 0.25 % of the whole study area.

The NDVI standard deviation of the vegetation period was found to be particularly useful. The initially tested period was April to October, which turned out to give many false positives due to remaining snow cover or delay in spring vegetation in April or early cold periods leading to discoloration of leaves in October. Still, with this approach, misclassifications occurred, as the Landsat time series is not so dense (repeat rate = 16 days). This means, if the data of satellite overpass, there is clouds or rain, no image is acquired for a month. After one month, the spectral response to the typical human usage such as mowing or tilling is lost and can thus not be properly detected. Therefore, we included a final post-processing step using Copernicus Sentinel-2 SR data from the past two years, which have meanwhile become available in GEE. It applies as well for Sentinel-2 data that clouds need to be removed to get rid of wrong reflectance values. Also, in this case a pre-defined GGE algorithm, which uses the additional “QA_60” band (cloud mask) that is available for each image, was used to cut out clouds. The standard deviation of the NDVI over the past two years was calculated from the S2 data. All areas with a maximum standard deviation exceeding 0.17 were masked from the previous result. Figure 14 shows, how different land cover types appear in a Landsat 8 time series. Standard deviation is clearly higher for annual crops than for pastures. This huge mapping effort was done for identification of underutilized areas throughout Europe and the neighboring Ukraine. In the last step, the masking of specific areas was performed, see next section.



(a) *(b)*

Figure 14: NDVI Time Series for used (a) and underutilized land (b) for the last five years. Same areas shown in Google Earth (a) and Google Street View (b)

2.4 Masking Specific Areas

In the frame of BIOPLAT-EU, it was decided to exclude specific areas from all further consideration. These so-called “cut-out areas” comprise:

1) **Forest areas** (HRL Forest)
   Forest areas are removed from potential MUC lands, as the STEN tool is not considering land cover change from forest to energy crops.
2) **Settlement area** (HRL Imperviousness, Open Street Map (OSM) and CORINE land cover)
   Some settlement areas with their specific mixture of sealed areas, gardens, trees and shrubs can sometimes be spectrally similar to underutilized areas due to the spatial resolution of 30 m as used in TIER-1. Therefore, these areas are also removed to avoid potential commission errors.
3) **Water and Wetland areas** (HRL Water & Wetlands)
   Water and wetland areas are also removed due to limitations in drivability for mechanized growing of bioenergy crops.
4) **Protected areas** (Natura2000)
   Protected areas are removed totally, although the consortium is aware, that crops used for energy might be allowed in some protected areas (e.g. outer zones of national parks). However, due to missing European-wide spatial separation between allowed and restricted zones, we removed all areas to avoid critical land competition.
5) **Steep slopes** (> 15 ° slope in Shuttle Radar Topography Mission digital elevation model (SRTM))
   Steep slopes with inclinations larger than 15° are also removed because for these areas, mechanized land cultivation is typically not feasible.
6) **Other not usable areas** (CORINE land cover)
   Other not usable areas like beaches, bare rocks or glaciers (CLC classes 331, 332 and 335) are also eliminated.
7) **Agriculturally used areas only for underutilized land** (CORINE land cover)
   From the agriculturally used areas, most classes (CLC classes 221, 222, 223, 231, 241, 242 and 244) are only removed from the category “underutilized land”, not from the “contaminated land”. This is due to the fact, that we consider contaminated land shall not be used for food production. If they are nonetheless presently used for food production, we can still consider them as potential land for energy crops. Please note, that point 6 and 7 from the list above are covered in one-step, as the additional data set for the cut-out is the same (CORINE land cover). The annual crops in CORINE land cover (CLC classes 211, 212 and 213) are not removed from the underutilized areas per se in order to detect abandoned farmlands. Abandoned permanent crops cannot be detected, because the spectral signatures are too similar to active permanent crops.

In the last step, also a minimum mapping unit is applied. For TIER-1, 10 ha is the minimum mapping unit (MMU). The whole masking procedure is depicted schematically in Figure 15.
All of the above described steps are valid for the EU and not for Ukraine, as many of the data sets are not available for Ukraine. Nonetheless, we used a similar processing employing national data of land use (for forest, settlement, water, other and agriculture) and protected areas. Further, an agricultural map of the Sen2Agri project was employed for cutting out annual crops. OSM data and slope map were equivalent to the inputs for the EU.

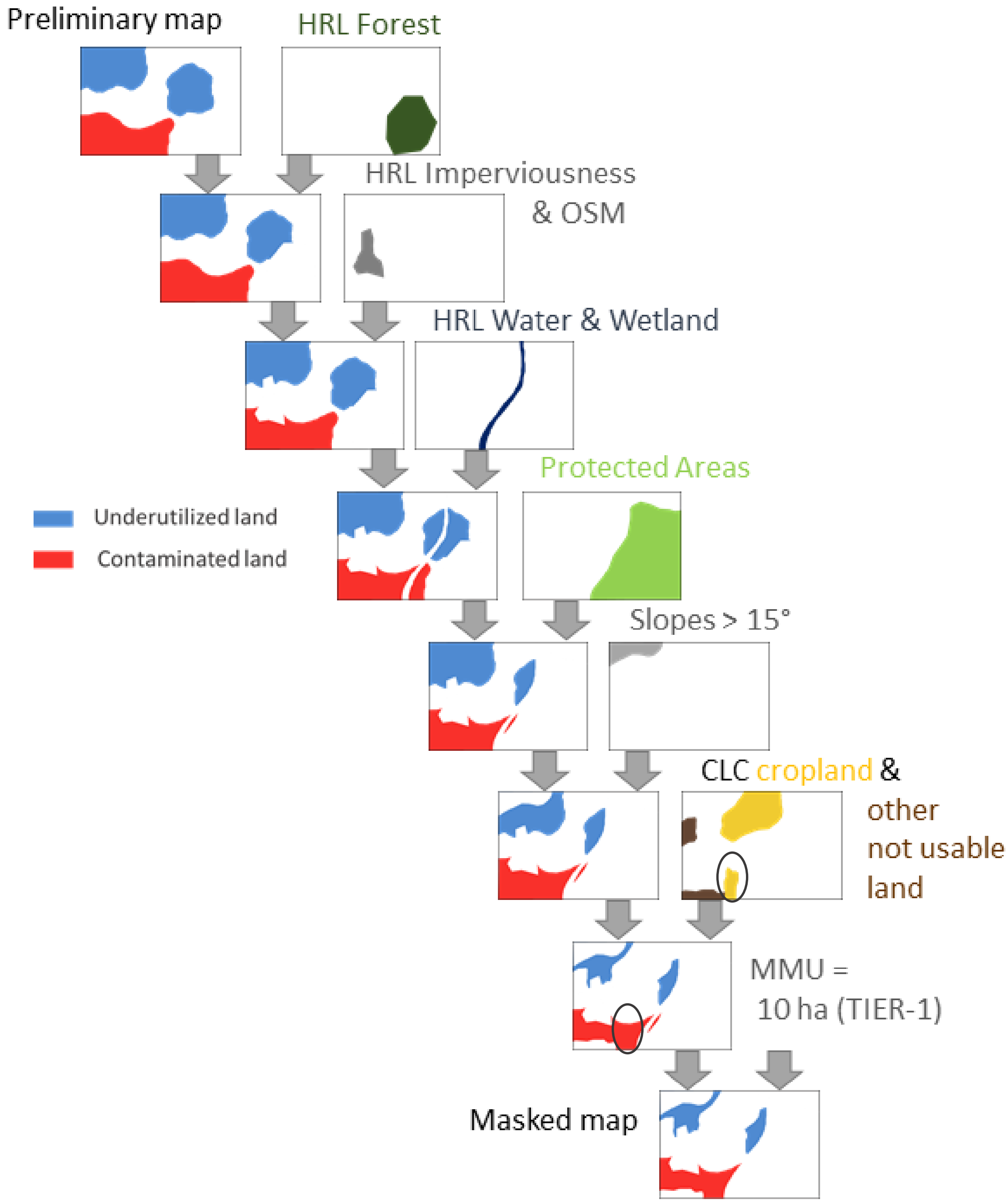


Figure 15: Schematically procedure of cutting out areas from preliminary results

The final shape of the masked map is thus determined by the data used for masking. This fact should be kept in mind, when evaluating the delineation of the MUC lands. An example of such an effect is shown in the Results (Chapter 3). Figure 21 shows a well-classified underutilized land in Italy, which is bordered by neighboring steep slopes determining the outline of the area.

2.5 Post-processing for integration in the WebGIS

Although the result after the cutting out and applying the MMU can be considered final, there are some additional post-processing steps, which are needed for the proper integration into the WebGIS service. These post-processing steps include:

1. Smoothing of the border for a smoother appearance
2. Simplification of the polygon structures to reduce the weight of the data (reduction of vertices per polygon) and thus allowing easier manipulation in the WebGIS
3. Re-assessment and application of the MMU (if needed)
4. Calculation of a set of attributes for the interaction with STEN (see Table 3 for details).

The effect of steps 1 and 2 are shown in Figure 16 b) and c). Figure 16 a) depicts the result of the classification after refinement. Since it is a pixel-based approach, the individual pixels are still visible. By applying the before mentioned post-processing steps, the results is more intuitive and has a logical appearance for the user of the WebGIS. One example of an underutilized area in Italy before and after the post-processing is shown in Figure 17.

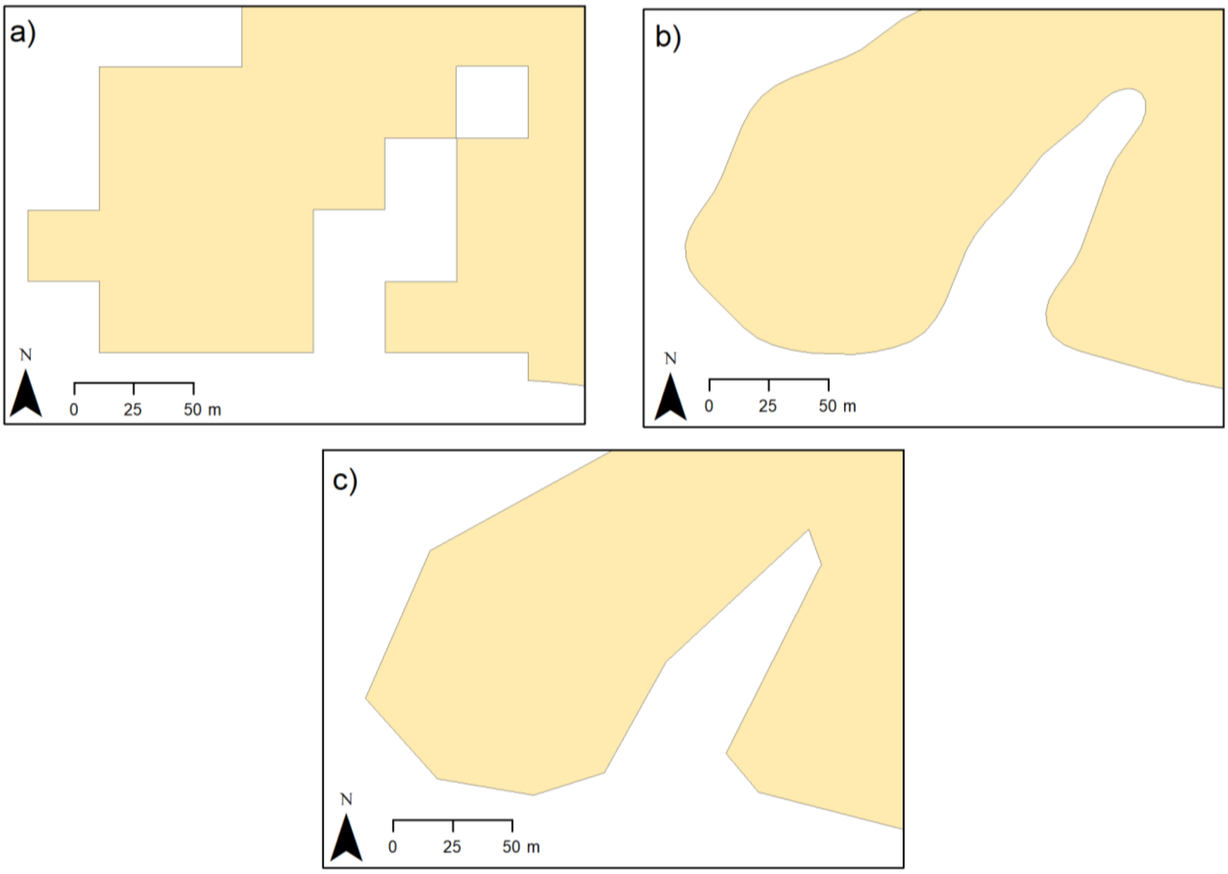


Figure 16: Schematic example of results of individual post-processing steps: a) result of masking specific LU categories; b) result of smoothing algorithm; c) result of simplify algorithm

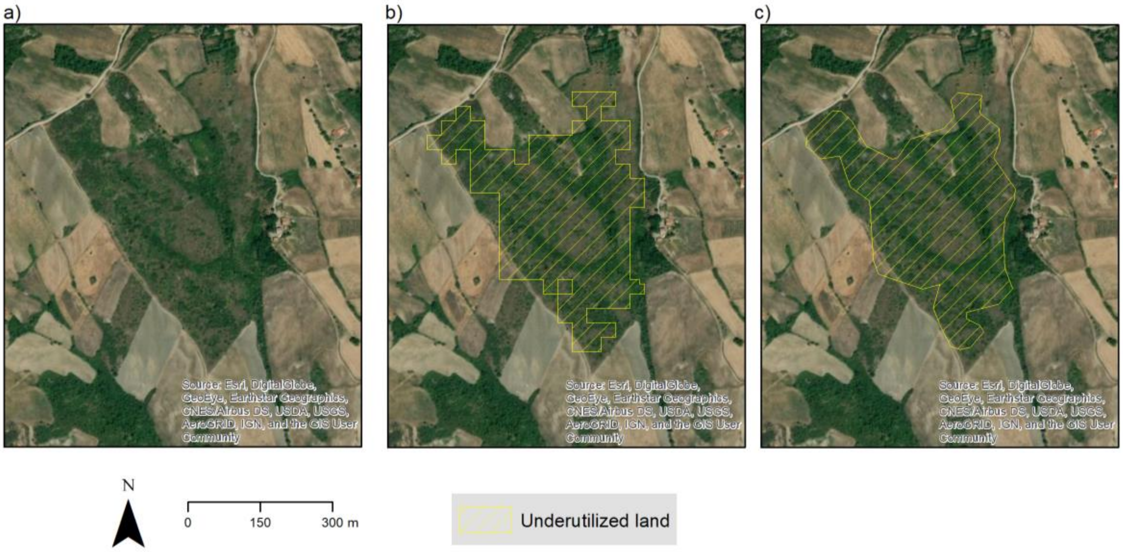


Figure 17: Example of underutilized area in Italy: a) VHR image from the ArcGIS basemap; b) raw result c) final post-processed result

Step 4 is needed in order to provide the STEN tool with needed input data. The list of attributes (Table 3) has been developed together with FAO and UCLM. Two separate layers are produced: underutilized land with only the attributes given in the first part of Table 3 and contaminated lands with all attributes given in Table 3. The attributes are delivered as a separate table to the shapefile with the MUC_ID being the unique identifier to link the map with the table.
The sources of the attributes are mainly from EUROSTAT for Europe and national statistics for Ukraine (LAU_CODE, NUTS3_CODE, CNTR_CODE). NAME_CODE is derived from these attributes and MUC_ID is a unique identifier. The additional attributes for contaminated lands are needed in order to calculate the land cover change effect within STEN. Data source for all of these attributes are CORINE land cover for Europe and the national land cover classification for Ukraine. They are the same sources, which are also used in the generation of the Target Area Base Layer (TABL), which was generated in the frame of WP3 and will be explained in the respective deliverable.

Table 3: Attributes for the underutilized and contaminated lands layer

FIELDS	DESCRIPTION
MUC_ID	Identifier unique
NAME_CODE	NUTS3_CODE + LAU_CODE + MUC_ID
CNTR_CODE	ISO 3166-1 – Alpha-2-code
NUTS3_CODE	Nomenclature of Territorial Units for Statistics (ASSIGN MUC CENTROID)
LAU_CODE	Local Administrative Units (ASSIGN MUC CENTROID)
TYPE	Type: U, C
AREA	Area
Separate and additional table for contaminated lands only with the following attributes:	
MUC_ID	Identifier unique
AnCro	Share of annual crops
PeCro	Share of permanent crops
Past	Share of pastures
Oth	Share of other

3 Results

3.1 Contaminated lands

The results of the contaminated land mapping for whole Europe with the national thresholds based on the Heavy Metal contamination in soils is shown in Figure 18. For most countries, national thresholds were used. Only for two countries (Slovenia, Bulgaria), where no national thresholds could be found, the relevant EU thresholds were applied. It has to be mentioned, that areas of contaminated land contain agriculturally used lands, as this was not cut out during the post-processing (for reasoning please see Point 7 in Chapter 2.4). Since this layer is based on an existing product, the quality can only be evaluated based on the quality of the input data. In addition, accuracy assessment of contaminations would involve high efforts for in-situ soil sampling and lab diagnostics, which have not been foreseen or possible in BIOPLAT-EU. Therefore, no specific accuracy assessment can be done for the contaminated land layer.

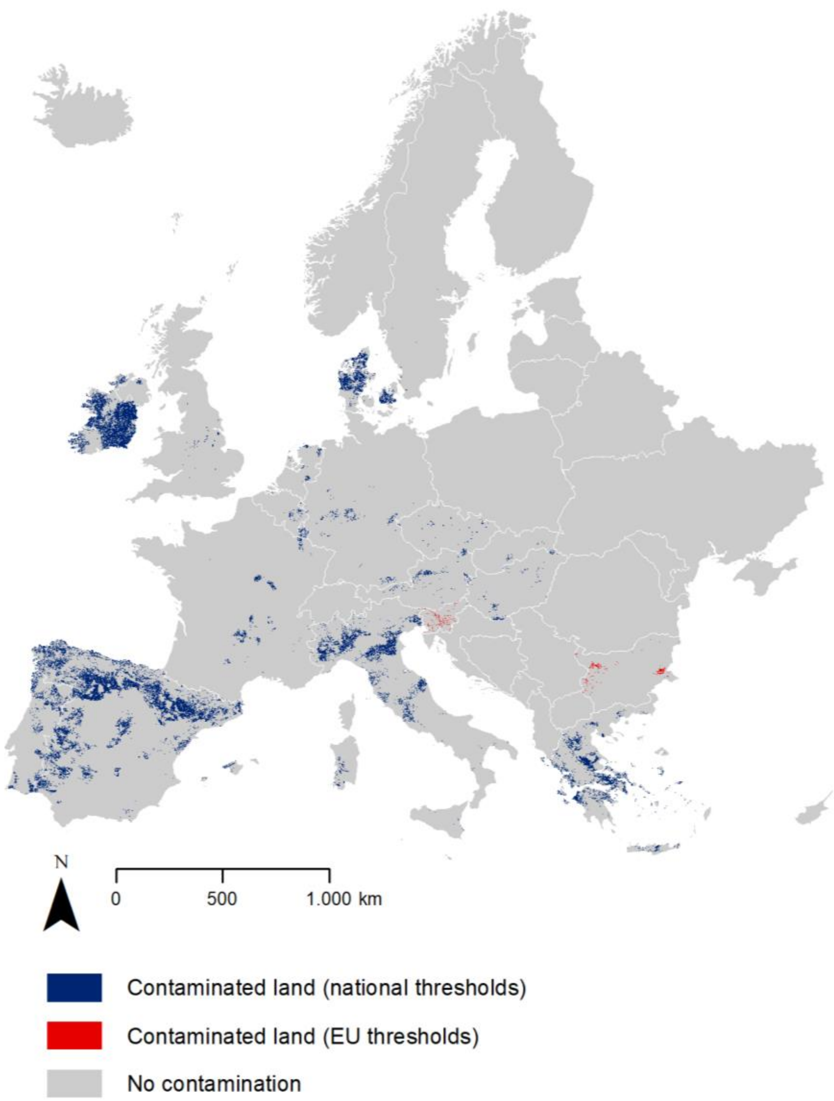


Figure 18: Map of contaminated soils in Europe using the national thresholds. For countries lacking national thresholds, the relevant EU thresholds are applied.

National maps of contaminated land can be very different from the result of the given approach. There are multiple reasons for this difference. One is the difference in ground data (other than the LUCAS sample points), another one is the higher resolution of the interpolation (JRC map 1 x 1 km) and finally, the pollutants can also be different. These differences can be seen in Italy in Figure 19: while large areas are considered contaminated using the JRC map, the national data set is much more selective.

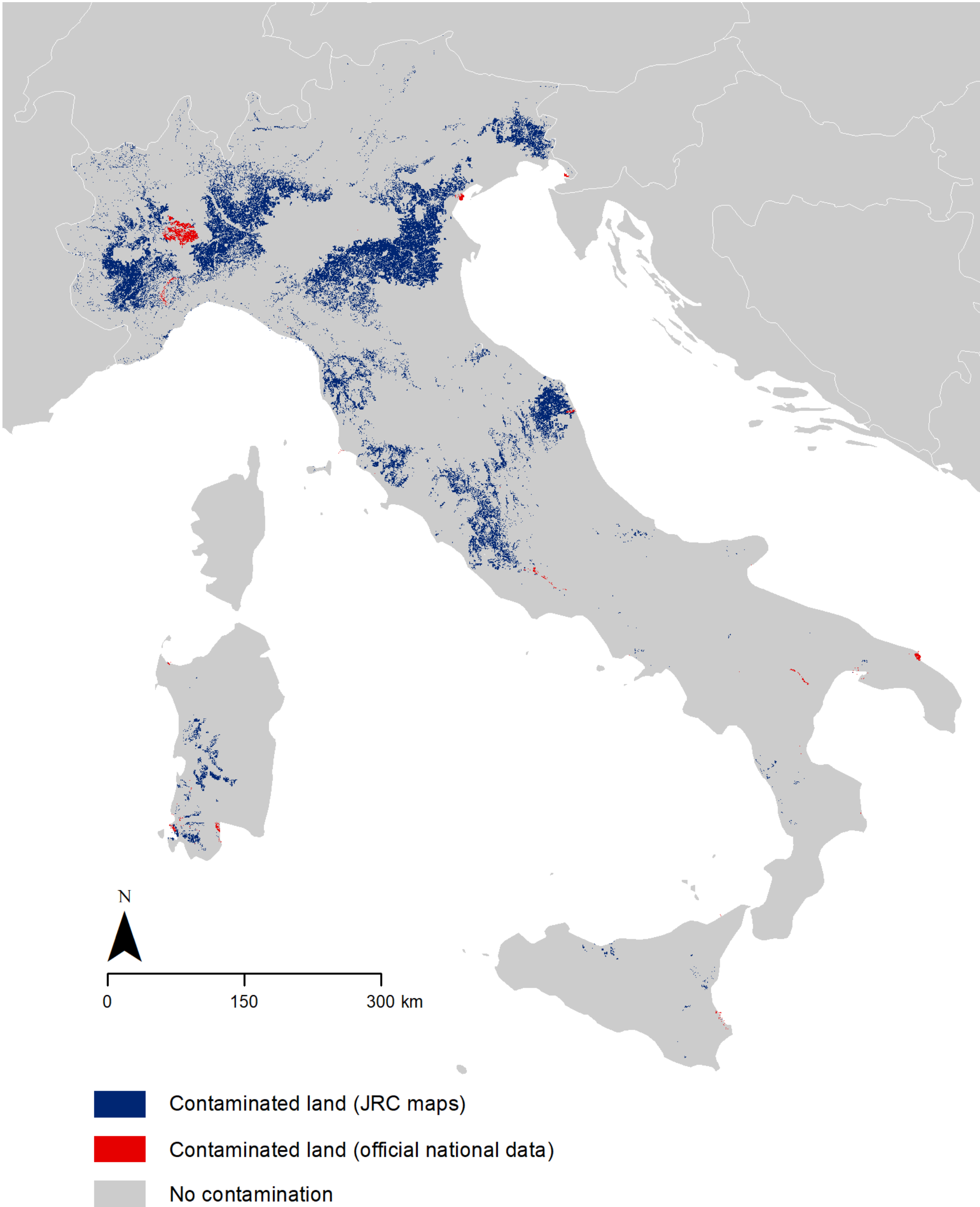


Figure 19: Italian map of contaminated soils from JRC data and national sources

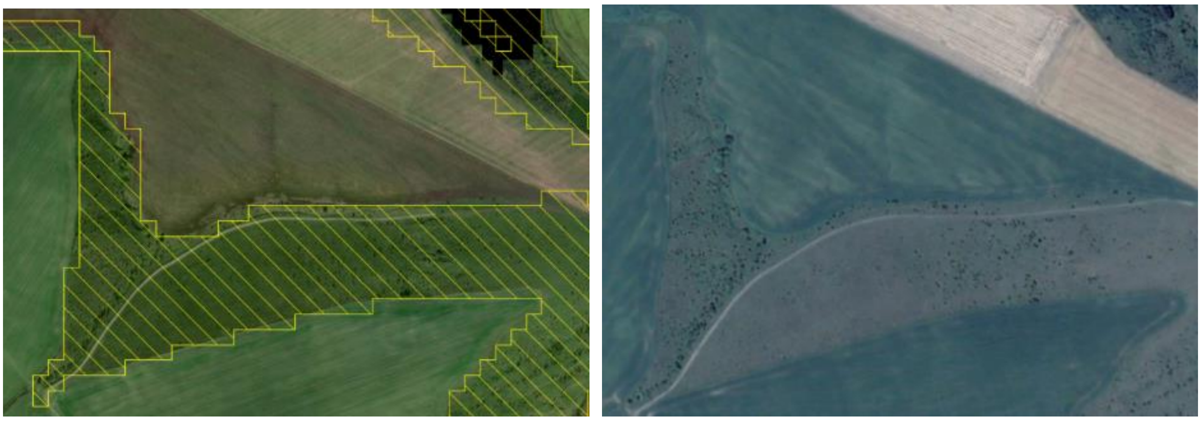
3.2 Underutilized Lands

The extent of underutilized lands is strongly varying in Europe. There are obvious factors such as the economic situation, soil and climate conditions, historic development etc. Some examples of correctly classified underutilized lands are shown in Figure 20. The underutilized areas are clearly separated from agriculturally used areas or from forest.

Figure 21 shows the effect of the application of cut-out masks, in this specific case the mask of steep slopes. Neighbouring areas look similar to the identified underutilized land, but are not included in the mask. This is due to the inclination of the slope there. The difference in the outline is due to the smoothing algorithm in the post-processing step.

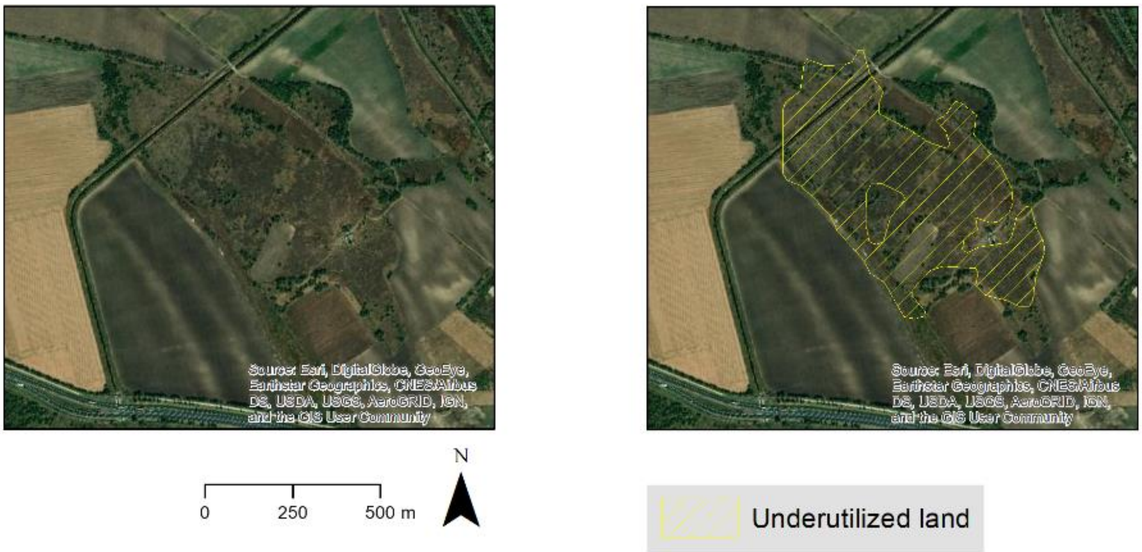
At this point, the issue of the used background image needs to be raised. The background image is the image, which the map is compared to (in the examples below but also in the final WebGIS tool). Google Earth or ArcGIS base map, which are commonly used as background and thus for comparison, are snapshots of a specific day – the day, when the respective image was taken. This image acquisition time can sometimes be more than five years ago or the image could have been taken during winter, when the land cover looks very different. This make comparison sometimes difficult. In addition, many more details are visible in these images due to their very high spatial resolution (usually better than 1 m) compared to the used input images of the map (30 m). Clearly, all these factors have to be taken into account, when comparing the result of a time series of 30 m spatial resolution with such a background image.

A proper accuracy assessment will be performed on the final version of the underutilized TIER-1 map by using stratified sampling and visual interpretation of sample points using Google Earth temporal view. This exercise will only be done after some minor corrections foreseen after reception of feedback on the preliminary map from the project partners. First accuracy checks on individual bio-geographical regions - with limited area sampling and no area-specific weights applied - indicate high overall accuracies of around 80%. User's accuracy of underutilized class ("is the mapped area really underutilized?") is typically a bit lower (around 75%). In the minor adjustment by visual checks, we consider to increase the level of user's accuracy of underutilized class by removing wrong polygons. All accuracy assessments and the final maps are planned to be published in the frame of BIOPLAT-EU project.



(a)

(b)



(c)

Figure 20: Examples of well-classified underutilized land in (a) Ukraine, (b) Italy and (c) Hungary

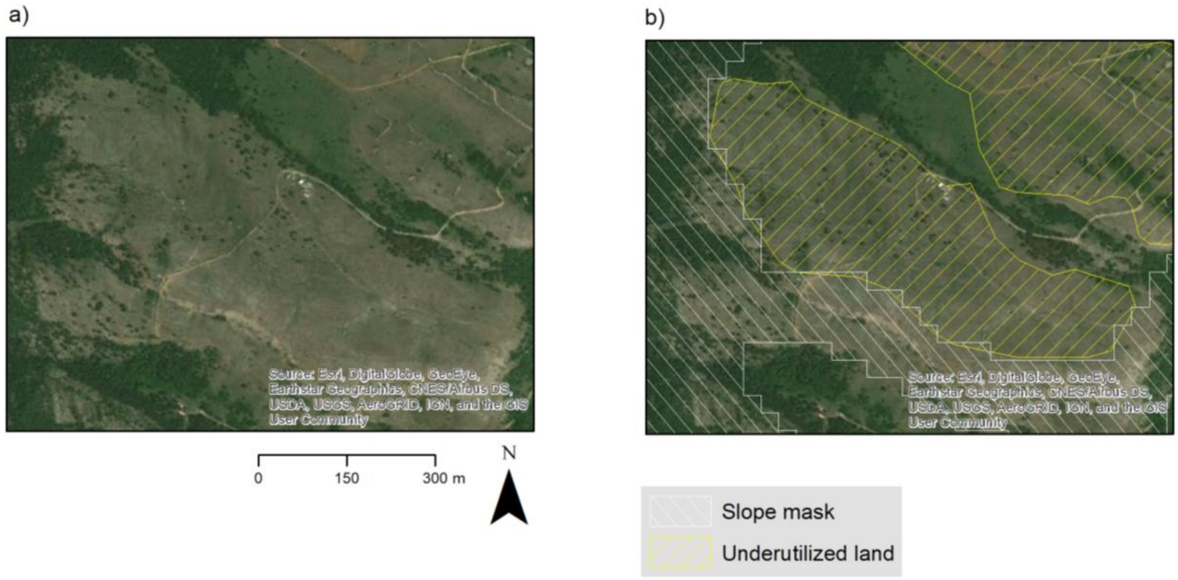


Figure 21: Example of well-classified underutilized land in Italy with the separation from neighboring steep slopes

Unfortunately, not all areas are classified as accurately as the examples above. There are several sources of errors and misclassifications, which also have to be mentioned.
Firstly, cleared forest areas are sometimes misclassified as underutilized land. These areas do not show any obvious signs of human intervention for five years or more. The trees are planted and then they are usually left to grow without intervention. In addition, these areas are usually not part of the high resolution Copernicus "tree cover density" or "forest type" layer, as they are not yet fully covered by trees at the time, when the Copernicus layer was produced. Therefore, such clear cuts and young forests/regeneration areas represent errors in the classification (see Figure 22 left). Similarly, short rotation coppice plantations are sometimes classified as underutilized in the first years after being planted for the same reasons as the forest gaps. The latter happens most frequently in Hungary, an example is shown in Figure 22 right.

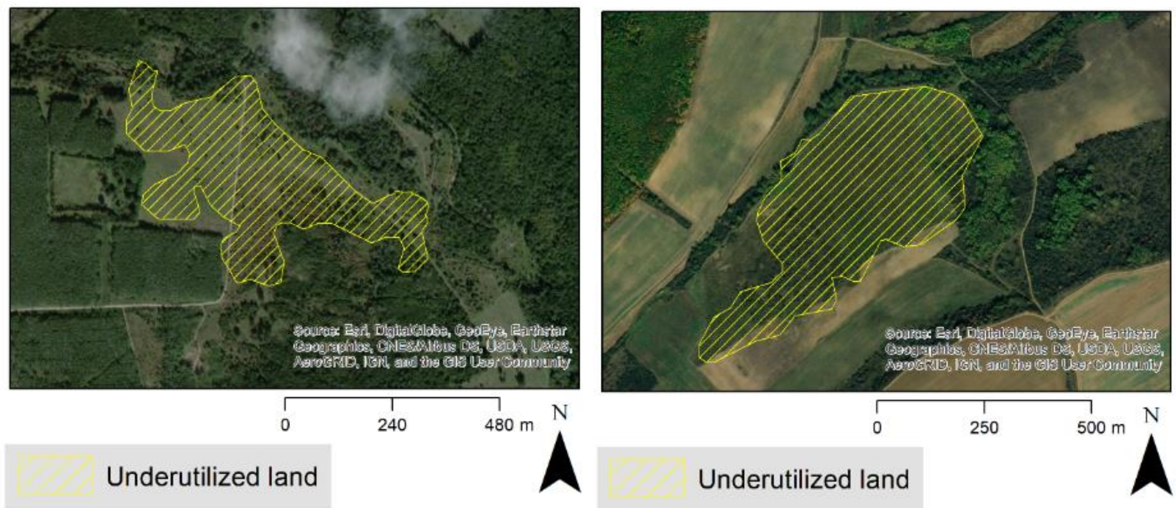


Figure 22: Example of misclassified underutilized lands in forest gaps (left) or on short rotation coppice area (right), both in Hungary.

The second frequent misclassification occurs in areas with very narrow strips of agricultural fields (typically less than 30 m wide). Since the spatial resolution of the input data is 30 m, smaller signatures cannot be resolved and can therefore only be represented as mixed pixels. Due to this mixed signal, the individual human interventions at the individual stripes become blurred and the whole area often appears like underutilized land. This effect has been observed in Ukraine or in Romania, an example of this situation is shown in Figure 23.

Figure 23: Example of misclassified underutilized lands due to small agricultural stripes in Romania.

The third source of misclassifications are permanent crops, which are not properly covered by CORINE land cover, mainly due to limited size. Permanent crops, especially olive groves or extensive fruit yards, show a very similar spectral and temporal behaviour as underutilized lands with scattered trees. Sometimes, they are not even distinguishable in the very high resolution data from Google Earth. Figure 24 shows such an area, where the differentiation between olive groves and areas with other scattered trees is only possible based on the Google Street View imagery. Thus, in these cases, misclassifications in the 30 m data are possible.

A full European map of underutilized areas is difficult to present, as the patches are rather small. Therefore, we show selected map of Spain exemplarily in Figure 25.

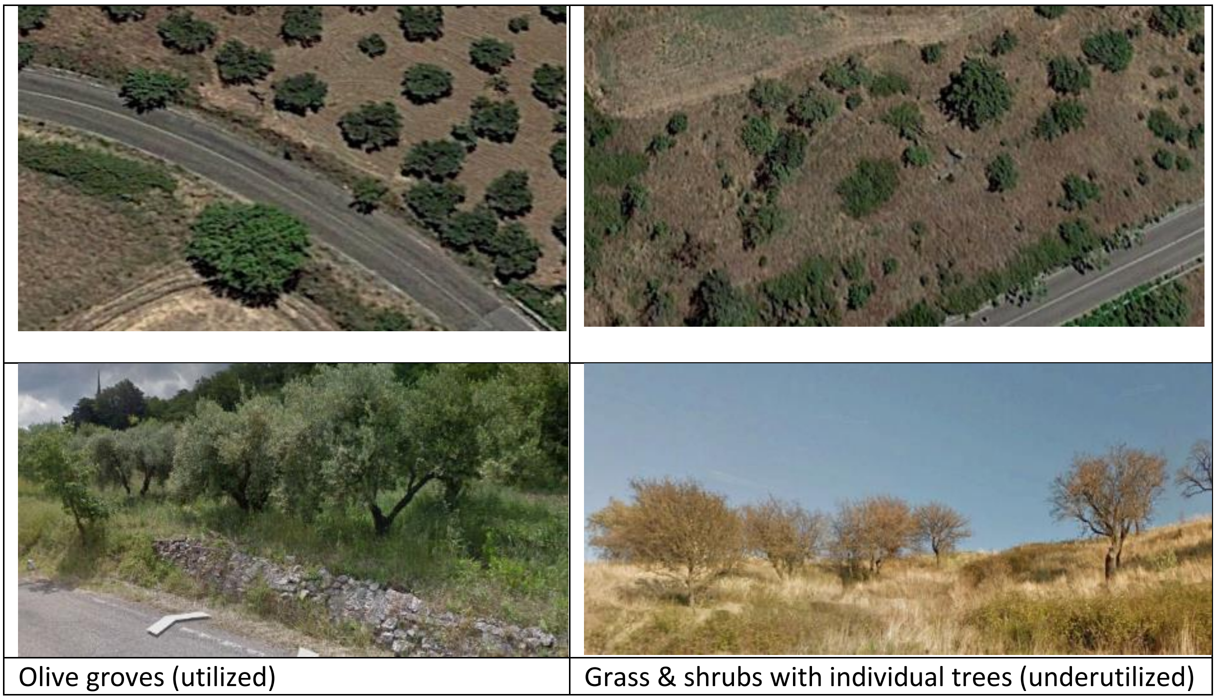


Figure 24: Small patches of permanent crops (mainly olive groves) with similar spectral and temporal appearance as abandoned or underutilized lands with scattered trees

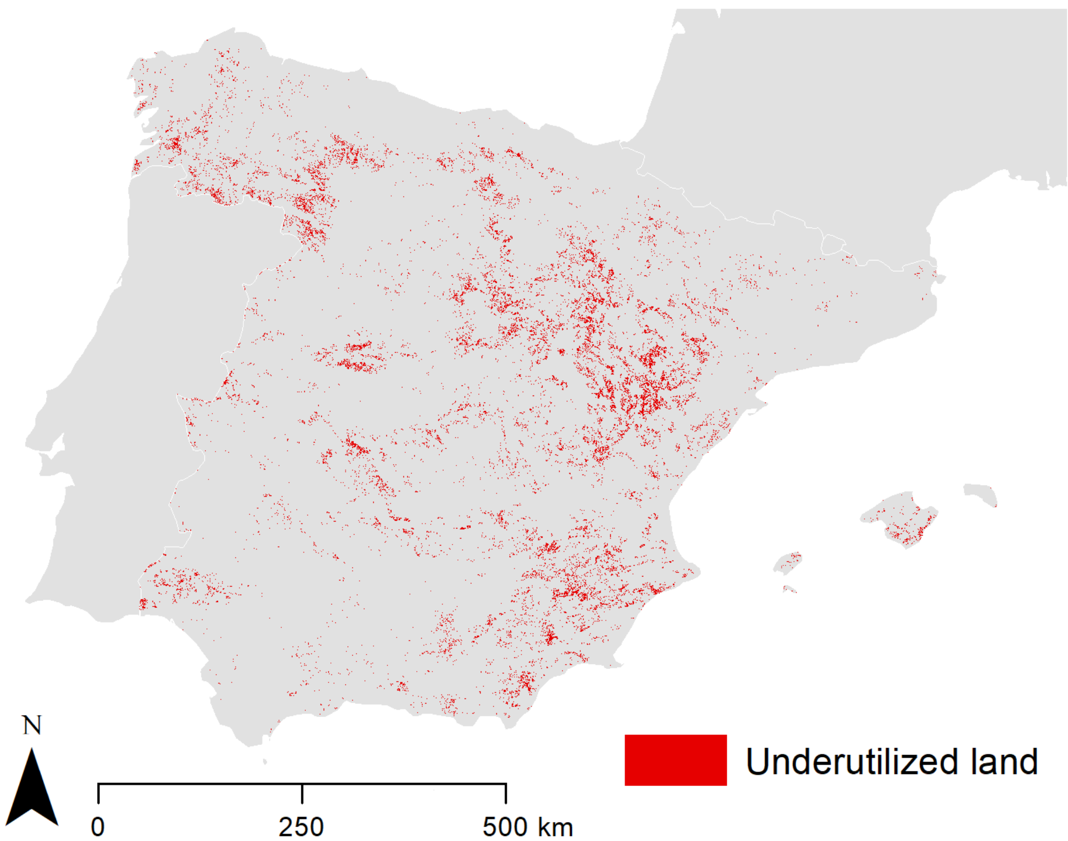


Figure 25: Preliminary TIER-1 map of underutilized land for Spain

4 Conclusions

This deliverable provided the description of all input data and processing methods for the generation of the TIER-1 map of Europe consisting of contaminated and underutilized lands. The marginal lands were discarded due to the contradiction with BIOPLAT-EU's definition to avoid the "food versus fuel" problem. Contaminated lands were identified based on the European-wide map of heavy metals in soils produced by the Joint Research Centre (JRC). On this heavy metal contamination map, we applied two different sets of thresholds. The first set are common thresholds for Europe used in publications in the past. The second set consists of national thresholds. The national thresholds were finally used for the generation of the map, as their results were considered more reliable and plausible by the local project partners and better in line with national statistics.

The mapping of underutilized land was based on a remote sensing time series approach using Landsat data with a spatial resolution of 30 m instead of the originally in the proposal foreseen MODIS data with 250 m spatial resolution. The classification was done separately for each bio-geographical region of Europe, as the typical appearance and signatures of underutilized land, but also for the different utilized land classes (forest, agriculture, pasture, etc.), vary considerably. Training data was generated for each region and classification was performed. In a next step, areas not usable for bioenergy production due to different reasons were cut out and a post-processing was performed to allow the data to be suitable for the WebGIS integration in Work Package 3. Finally, the current version of the map is considered preliminary, as the project partners and selected users can review, comment and improve this version, before the final map is provided. This final version of the underutilized land map will then also undergo restrictive accuracy assessment by stratified random sampling in addition to the quality checks performed to date.

5 References

Alcantara, C.; Kuemmerle, T.; Baumann, M.; Bragina, E. V.; Griffiths, P.; Hostert, P.; Knorn, J.; Müller, D.; Prishchepov, A. V.; Schierhorn, F.; Sieber, A. & Radeloff, V. C. (2013). Mapping the extent of abandoned farmland in Central and Eastern Europe using MODIS time series satellite data. Environmental Research Letters, 2013, 8.

Ali, J.; Khan, R., Ahmad, N.; and Maqsood, I. (2012). Random forests and decision trees. International Journal of Computer Science Issues (IJCSI), 9(5).

Amlinger, F.; Pollak, M.; Favoino, E. (2004). Heavy Metals and Organic Compounds from Wastes Used as Organic Fertilisers. Final Report ENV.A.2./ETU/2001/0024. 244 p. Pp. 73-74.

Bundesministerium für Justiz und Verbraucherschutz, (1998). Bundes-Bodenschutzgesetz vom 17. März 1998 (BGBl. I S. 502), das zuletzt durch Artikel 3 Absatz 3 der Verordnung vom 27. September 2017 (BGBl. I S. 3465) geändert worden ist. https://www.gesetze-im-internet.de/bbodschg/BBodSchG.pdf

Bundesministerium für Justiz und Verbraucherschutz, (1999). Bundes-Bodenschutz- und Altlastenverordnung vom 12. Juli 1999 (BGBl. I S. 1554), die zuletzt durch Artikel 3 Absatz 4 der Verordnung vom 27. September 2017 (BGBl. I S. 3465) geändert worden ist. https://www.gesetze-im-internet.de/bbodschv/BBodSchV.pdf

Bundesministerium für Justiz und Verbraucherschutz, 1981: Futtermittelverordnung in der Fassung der Bekanntmachung vom 29. August 2016 (BGBl. I S. 2004), die zuletzt durch Artikel 2 der Verordnung vom 18. Juli 2018 (BGBl. I S. 1219) geändert worden ist. http://www.gesetze-im-internet.de/futtmv_1981/FuttMV_1981.pdf

LUBW (Landesanstalt für Umwelt Baden-Württemberg (Hrsg.)), 2018: Bodenschutzrecht. Handreichung für die Verwaltung. Landesanstalt für Umwelt Baden-Württemberg. Karlsruhe, 480 S. online verfügbar unter: https://fachdokumente.lubw.baden-wuerttemberg.de/servlet/is/102234/Fachzugang_Recht_Internet.pdf?command=downloadContent&filename=Fachzugang_Recht_Internet.pdf&

Bundesministerium für Justiz und Verbraucherschutz, 2010: Kontaminanten-Verordnung vom 19. März 2010 (BGBl. I S. 286, 287), die zuletzt durch Artikel 2 der Verordnung vom 24. November 2016 (BGBl. I S. 2656) geändert worden ist. https://www.bmel.de/SharedDocs/Downloads/DE/_Verbraucherschutz/Lebensmittelsicherheit/Kontaminanten-Verordnung.pdf?__blob=publicationFile&v=2

Brooks, R.; Morrison, R.; Reeves R.; Dudley, T. & Akman, Y. (1979). Hyperaccumulation of Nickel by Alyssum Linnaeus (Cruciferae). Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain). 203. 387-403. 10.1098/rspb.1979.0005.

Carlon, C. (Ed.) (2007). Derivation methods of soil screening values in Europe. A review and evaluation of national procedures towards harmonization. European Commission, Joint Research Centre, Ispra, EUR 22805-EN, 306 pp.

Chaney, R.; Angle, J.; Broadhurst, C.; Peters, C.; Tappero, R. & Sparks, D. (2007). Improved Understanding of Hyperaccumulation Yields Commercial Phytoextraction and Phytomining Technologies. Journal of environmental quality. 36. 1429-43. 10.2134/jeq2006.0514.

Colditz, R. R. (2015). An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. Remote Sensing, 7(8):9655–9681.

Council of the European Union, (2002). Directive 2002/32/EC of the European Parliament and of the Council of 7 May 2002 on undesirable substances in animal feed - Council statement. Official

Journal L 140, 30/05/2002 P. 0010 - 0022. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1591166382702&uri=CELEX%3A32002L0032

Dalponte, M.; Ørka, H. O.; Gobakken, T.; Gianelle, D. & Næsset, E. (2013). Tree species classification in boreal forests with hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 51(5):2632–2645.

Estel, S.; Kuemmerle, T.; Alcántara, C.; Levers, C.; Prishchepov, A. & Hostert, P. (2015). Mapping farmland abandonment and recultivation across Europe using MODIS NDVI time series. Remote Sensing of Environment, Elsevier BV, 2015, 163, 312-325.

Estel, S.; Kuemmerle, T.; Levers, C.; Baumann, M. & Hostert, P. (2016). Mapping cropland-use intensity across Europe using MODIS NDVI time series. Environmental Research Letters, IOP Publishing, 2016 , 11, 024015.

European Commission, 2006: Commission Regulation (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1591166803328&uri=CELEX:32006R1881

Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Hughes, M.J. & Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. In: Remote Sensing of Environment, 194, 379-390.

Friedl, (2019). Derivation of glaciological parameters from time series of multi-mission remote sensing data - Applications to glaciers in Antarctica and the Karakoram. PhD Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 244 p. available at: https://www.researchgate.net/publication/337973916_Derivation_of_glaciological_parameters_from_time_series_of_multi-mission_remote_sensing_data_-_Applications_to_glaciers_in_Antarctica_and_the_Karakoram/figures?lo=1

Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D. & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 2017, 202, 18 - 27

Grinand, C.; Rakotomalala, F.; Gond, V.; Vaudry, R.: Bernoux, M. & Vieilledent, G. (2013). Estimating deforestation in tropical humid and dry forests in madagascar from 2000 to 2010 using multi-date landsat satellite images and the random forests classifier. Remote Sensing of Environment, 139:68–80.

Horning, N. et al. (2010). Random forests: An algorithm for image classification and generation of continuous fields data sets. In Proceedings of the International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences, Osaka, Japan, volume 911.

Li, T.; Ni, B.; Wu, X.; Gao, Q.; Li, Q. & Sun, D. (2016). On random hyper-class random forest for visual classification. Neurocomputing, 172:281–289.

Liaw, A. & Wiener, M. (2002). Classification and regression by random forest. R news, 2(3):18–22.

Löw, F.; Prishchepov, A.; Waldner, F.; Dubovyk, O.; Akramkhanov, A.; Biradar, C. & Lamers, J. (2018). Mapping Cropland Abandonment in the Aral Sea Basin with MODIS Time Series. Remote Sensing, MDPI AG, 2018, 10, 159.

Mellor, A.; Boukir, S.; Haywood, A. & Jones, S. (2015). Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. ISPRS Journal of Photogrammetry and Remote Sensing, 105:155–168.

Moldoveanu, A. M., (2014). Assessment of Soil Pollution with Heavy Metals in Romania. In: Hernandez-Soriano, M.C. (Ed.), 2014, Environmental Risk Assessment of Soil Contamination, https://www.intechopen.com/books/environmental-risk-assessment-of-soil-contaminationassessment-of-soil-pollution-with-heavy-metals-in-romania

Paz-Ferreiro, J.; Lu, H.; Fu, S.; Méndez, A. & Gascó, G. (2014). Use of phytoremediation and biochar to remediate heavy metal polluted soils: a review. Solid Earth. 5. 10.5194/se-5-65-2014.

Sallustio, L.; Pettenella, D.; Merlini, P.; Romano, R.; Salvati, L.; Marchetti, M. & Corona, P. (2018). Assessing the economic marginality of agricultural lands in Italy to support land use planning. *Land Use Policy*, 76:526–534.

Soukup, T.; Brodský, L. & Vobora, V. (2009). REAL: Remote sensing identification and monitoring of abandoned land, GISAT, 2009. Available at: http://www.gisat.cz/images/upload/3bd38_eenvi-gisat-real.pdf.

Sipos, P.; & Poka, T. (2002). Threshold Limit Values for Heavy Metals in Soils in the Function of Spatial and Temporal Variation of Geochemical Factors. https://www.researchgate.net/publication/253733176 , 8p.

Szatmári, D.; Kopecká, M.; Feranec, J. & Goga, T. (2018). Abandoned Agricultural Land Mapping using SENTINEL-2A Data. Proceedings of 7th International Conference on Cartography and GIS, Sozopol, Bulgaria ISSN: 1314-0604, Eds: Bandrova T., Konečný M, 2018, p. 792-800.

Tóth, G., Hermann, T.; Silva, M. D. & Montanarella, L. (2016). Heavy metals in agricultural soils of the European Union with implications for food safety. Environment International, 88:299–309.

UBA (Ed), (2001). Sechster Umweltkontrollbericht 2001. https://www.umweltbundesamt.at/umweltsituation/umweltkontrollbericht/ukb2001/.

Vacha, R. & Sanka, M. (2014). Assessment of limit values of risk elements and persistent organic pollutants. Pp. 191-197. in soil for Czech legislation. In: Plant Soil and Environment. May 2014.

Witter, E. (2009). Agricultural use of sewage sludge - Is there a need to revise the Swedish regulations pertaining to heavy metals? (https://scholar.google.at/scholar?q=Witter+Ernst+agricultural+use+of+sewage&hl=de&as_sdt=0&as_vis=1&oi=scholart; This document is a complement to and update of the unpublished report “Land application of sewage sludge-review of national and international reports on allowable heavy metals loads” by Lage Bringmark (2002-04-19, Naturvårdsverket)). 17 p. P. 5.