

|  |  |
| --- | --- |
| D2.3a | Report on Methodology development |
|  | |
| **MAIL**: Identifying Marginal Lands in Europe and strengthening their contribution potentialities in a CO2 sequestration strategy | |

|  |  |
| --- | --- |
| Project title | Identifying Marginal Lands in Europe and strengthening their contribution potentialities in a CO2 sequestration strategy |
| Call identifier | H2020 MSCA RISE 2018 |
| Project acronym | MAIL |
| Starting date | 01.01.2019 |
| End date | 31.31.2021 |
| Funding scheme | Marie Skłodowska-Curie |
| Contract no. | 823805 |
|  | |
| Deliverable no. | D2.3a |
| Document name | MAIL\_D2.3a\_01\_20191029.docx |
| Deliverable name | Report on Methodology development |
| Work Package | WP2 |
| Nature[[1]](#footnote-1) | R |
| Dissemination[[2]](#footnote-2) | PU |
| Editor |  |
| Authors | Rodrigo Gómez Conejo, Natalia Verde |
| Contributors |  |
| Date |  |

# MAIL Consortium

|  |  |
| --- | --- |
| Aristotle University of Thessaloniki (AUTH) Greece | Industrieanlagen Betriebsgesellschaft MBH (IABG) Germany |
| Gounaris N. – Kontos K. OE (HOMEOTECH) Greece | Centrum Badan Kosmicznych Polskiej Akademii Nauk (CBK PAN) Poland |
| Universitat Politecnica de Valencia (UPV) Spain | Fundacion Centro De Servicios Y Promocion FOrestral Y de su Industria De Castilla y Leon (CESEFOR) Spain |

# Abbreviations

|  |  |
| --- | --- |
| **Term** | **Explanation** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Contents

[MAIL Consortium 3](#_Toc30586337)

[Abbreviations 4](#_Toc30586338)

[Document description 7](#_Toc30586339)

[Executive Summary 8](#_Toc30586340)

[1. Introduction 9](#_Toc30586341)

[2. Definition of minimum marginal land parcel area 11](#_Toc30586342)

[3. Datasets used 11](#_Toc30586343)

[3.1 CORINE Land Cover 11](#_Toc30586344)

[3.2 Economic statistics 12](#_Toc30586345)

[3.3 Pan-European High Resolution Layers (HRL) 13](#_Toc30586346)

[3.4 Digital European elevation model 15](#_Toc30586347)

[3.5 Environmental dataset 16](#_Toc30586348)

[3.6 European Soil Database Derived data 17](#_Toc30586349)

[4. Definition of GIS specifications 18](#_Toc30586350)

[4.1 Coordinate Projection 19](#_Toc30586351)

[4.2 File format 19](#_Toc30586352)

[4.3 Coverage 19](#_Toc30586353)

[5. Review of existing research projects and literature for available GIS models & workflows 20](#_Toc30586354)

[5.1 Marginality index 20](#_Toc30586355)

[5.2 Land use change 21](#_Toc30586356)

[5.3 Economic data 21](#_Toc30586357)

[5.4 Tree Cover Density 22](#_Toc30586358)

[5.5 Digital Elevation Model of Europe 22](#_Toc30586359)

[5.6 Protected areas. European network of protected sites (Natura 2000) 22](#_Toc30586360)

[6. Development of the final ML detection, criteria and classification scheme. 23](#_Toc30586361)

[6.1 ML definition refinement 23](#_Toc30586362)

[6.2 Methodology overview 24](#_Toc30586363)

[6.3 Indicator / criteria selection 27](#_Toc30586364)

[6.4 Initial indicator classification categories and thresholds 29](#_Toc30586365)

[6.4.1 Land Use/Land Cover indicators 30](#_Toc30586376)

[6.4.2 Terrain and Soil indicators 30](#_Toc30586377)

[6.4.3 Sustainability indicators 34](#_Toc30586378)

[6.4.4 Productivity indicators 38](#_Toc30586379)

[6.4.5 Climate indicators 41](#_Toc30586380)

[6.5 ML classification schemes 43](#_Toc30586381)

[6.6 Production of intermediate layers 48](#_Toc30586382)

[7. Development and realization of a GIS using the datasets collected in T2.2 49](#_Toc30586383)

[7.1 Datasets overview 49](#_Toc30586384)

[7.2 Dataset transformation to common file formats, projection, etc 53](#_Toc30586385)

[7.3 GIS implementation 53](#_Toc30586386)

[7.4 Layers 55](#_Toc30586387)

[8. Evaluation of existing and development of indicator sets 56](#_Toc30586388)

[8.1 Selection of test sites 56](#_Toc30586389)

[8.2 Testing different sets of existing indicators and thresholds 56](#_Toc30586390)

[8.3 Development of new indicators sets 56](#_Toc30586391)

[8.4 Refinement of indicators and criteria 56](#_Toc30586392)

[9. Design and development of computation workflows. 56](#_Toc30586393)

[9.1 Workflow design for refined indicator sets 56](#_Toc30586394)

[9.2 Algorithm development for the production of new layers incorporating initial datasets 56](#_Toc30586395)

[10. Implementation of workflows to the GIS 56](#_Toc30586396)

[10.1 Workflow programming 56](#_Toc30586397)

[10.2 Automatization of procedures 56](#_Toc30586398)

[10.3 Workflow testing 56](#_Toc30586399)

[10.4 Produced tools description 56](#_Toc30586400)

[11. Development of guidelines for the usage indicators set and GIS workflows 56](#_Toc30586401)

[12. Experimental results for MLs identification 56](#_Toc30586402)

[13. 1st methodology refinement (indicator fine tuning based on the experimental results 56](#_Toc30586403)

[14. Refinement of results based on Τ2.4 results T2.8 recommendations 57](#_Toc30586404)

[References 58](#_Toc30586405)

[Annex I: Marginality index calculation and CLC reclassification 65](#_Toc30586406)

[Annex II: Table of figures 68](#_Toc30586407)

[Annex III: List of Tables 69](#_Toc30586408)

Do not forget to update the Contents

# Document description

**Document revision history**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Modifications introduced** | |
| **Modification reason** | **Modified by** |
| 0 | 17.10.2019 | First draft | X.XXXXXXXX |
|  |  |  |  |
|  |  |  |  |

(to be removed in the version for submission)

# Executive Summary

# Introduction

This document contents the methodology and key aspects to be taken into account for marginal land detection under the scope of MAIL project.

The starting point is the definition of marginal land as pointed out in Deliverable 2.1. As per aforementioned document marginality is driven by three main forces: environmental factors, socioeconomic factors and cultural factors

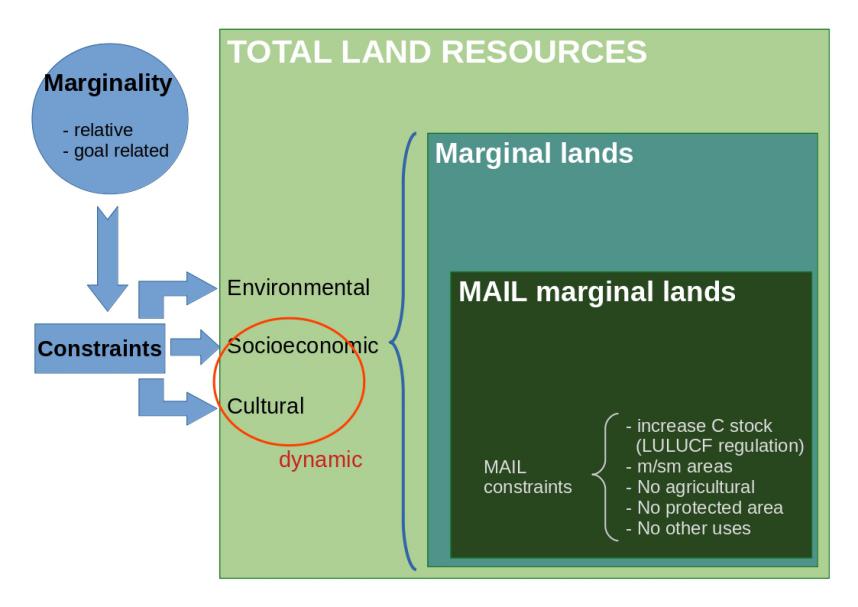


Figure 1: Transition between marginality and the definition on marginal lands in the framework of MAIL project.

These forces or factors directly influence land use, so that land cover data can be considered an aggregate of all these variables (Bertaglia et al. 2007). In addition, there are some constrains to relative marginality related with MAIL scope:

* Increase of C stock as per LULUCF regulation[[3]](#footnote-3)
* Mountainous and semi-mountainous areas
* No agricultural zones
* No protected areas
* No other uses

Marginal land detection was performed using primary data (forces related to the general definition of marginal lands) and secondary data (related to MAIL definition of marginal land). In the next tables are summarized the variables chosen for marginal land identification:

| **Factor** | **Dataset** | **Variable** |
| --- | --- | --- |
| Environmental factors | CORINE Land Cover | Land cover |
| Cultural factors |
| Socioeconomic factors | Economic statistics | Gross domestic product (GDP) |

Table 1: Factors for marginal land identification under **MAIL** scope. Source: Personal compilation

| **Constrain** | **Dataset** | **Variables** |
| --- | --- | --- |
| Increase of C stock as per LULUCF regulation | Pan-European High Resolution Layers (HRL) | Forest variables (Tree Cover Density) |
| Mountainous and semi-mountainous areas | Digital European Elevation Model | Elevation |
| No agricultural zones | CORINE land cover | Land cover |
| No protected areas | Environmental dataset | Protected areas |
| No other uses | CORINE land cover | Land cover[[4]](#footnote-4) |
| Other factors[[5]](#footnote-5) | European Soil Database Derived data | Soil variables |

Table 2: Constrains for marginal land identification under MAIL scope Source: Personal compilation

# Definition of minimum marginal land parcel area

Under MAIL project, the minimum area of marginal land[[6]](#footnote-6) was set (a priori) as 1 ha (10.000 m2). Marginal land with area lesser than this minimum parcel size will not be under the scope of MAIL project. Taking into account the European scale of the project, the aforementioned size was considered as equilibrium point between available resources and marginality characteristics at local scale.

# Datasets used

## CORINE Land Cover

In this study the CORINE Land Cover (CLC) database will be used as primary source of data. It provides a pan-European inventory of biophysical land-cover. The CLC inventory was initiated in 1985 (reference year 1990) and was created from remotely sensed data. Updates have been produced in 2000, 2006, 2012, and 2018. CLC provides information about land-cover changes for a substantial part of Europe. CLC uses a Minimum Mapping Unit (MMU) of 25 hectares for areal phenomena and a minimum width of 100 m for linear phenomena. The time series are complemented by change layers, which highlight changes in land cover with an MMU of 5 ha.

CORINE Land Cover database has been validated. The official classification accuracy of CORINE is ≥ 85% except for the release of 2002 (Büttner & Kosztra, 2017). However, there is no other accurate data available for the hole Europe and the data used for the analysis is considered appropriate for the spatial and temporal scale of the project.

CLC dataset will be utilized for general marginality detection and for the detection of land uses as agricultural and extensive grazing, that will compete with aforestation or other actions proposed under MAIL scope.

| **Specification** | **Source data specification** |
| --- | --- |
| File name: Corine Land Cover (CLC) 2018, Version 20 | Sensor: - |
| Coordinate system: ETRS89 LAEA | Data type: - |
| Production date: 14 - 06 - 2019 | Sensor resolution: - |
| Coverage (top L, BR coordinates): EEA 39 | Acquisition date: 2012 - 2018 |
| Grid size: 25 ha / 500 m | Grid size: - |
| Position accuracy: 100 m | Positional accuracy: - |
| Vertical accuracy: - | Vertical accuracy: - |
| Completeness: Complete | |
| File type, format: Vector, AutoCAD Slide (.sld) & ArcGIS Layer (.lyr) | |

Table 3: Technical specifications of CLC dataset. Source: Deliverable D2.2

## Economic statistics

As source of economic variables it was used the statistical data compiled by the European Statistical Office (EUROSTAT). EUROSTAT is the statistical office of the European Union, providing the European Union with statistics at European level that enable comparisons between countries and regions. Those statistics are related to Nomenclature of Territorial Units for Statistics or NUTS (French: Nomenclature des unites territoriales statistiques). NUTS is a three-level hierarchical geocode standard for referencing the subdivisions of countries for statistical purposes. The standard, adopted in 2003, is developed and regulated by the European Union, and thus only covers the member states of the EU in detail. The current NUTS 2016 classification is valid from 1 January 2018 and lists 104 regions at NUTS 1, 281 regions at NUTS 2 and 1348 regions at NUTS 3 level.

The variable chosen from EUROSTAT database at level 3, is Gross domestic product (GDP) at current market prices by NUTS 3 regions for the decade 2008 - 2017.

| **Specification** | **Source data specification** |
| --- | --- |
| File name: Gross domestic product (GDP) at current market prices by NUTS 3 regions | Sensor: - |
| Coordinate system: - | Data type: - |
| Production date: 1 - 08 - 2019 | Sensor resolution: - |
| Coverage (top L, BR coordinates): Europe | Acquisition date: 2008 - 2017 |
| Grid size: - | Grid size: - |
| Position accuracy: - | Positional accuracy: - |
| Vertical accuracy: - | Vertical accuracy: - |
| Completeness: Complete | |
| File type, format: Table, Microsoft Excel 97-2003 Worksheet (.xls) | |

Table 4: Technical specifications of EUROSTAT dataset. Source: Deliverable D2.2

## Pan-European High Resolution Layers (HRL)

Pan-European High Resolution Layers (HRL) provide information on specific land cover characteristics, and are complementary to land cover / land use mapping such as the CLC datasets. The HRLs are produced from satellite imagery through a combination of automated processing and interactive rule based classification. Since the production of the 2015 reference year the production is increasingly based on analyzing time series of satellite images from a number of different sensors, including the combination of optical and radar data. The main sources are the Sentinel Satellites (in particular Sentinel-2 and Sentinel-1). In addition to high resolution (HR) data, since 2015, very high resolution (VHR) imagery were also used for some of the products.

Five themes have been identified so far, corresponding with the main themes from CLC, i.e. the level of sealed soil (imperviousness), tree cover density and forest type, grasslands, wetness and water, and small woody features.

Tree Cover Density (TCD) will be utilized as forestry variable in order to select the marginal lands where activities such as aforestation and reforestation will have more impact on the C stock accounting system as defined by the LULUCF directive[[7]](#footnote-7). According this regulation, not all of the forest related sinks will count toward the mitigation target (Grassi et al., 2019), therefore carbon stock contribution, for legal purposes, will depend if performed on already managed forest lands or afforested and forested lands. Distinction between those categories will be performed through TCD thresholds in order to select the most suitable marginal lands for MAIL project.

| **Specification** | **Source data specification** |
| --- | --- |
| File name: Tree Cover Density (TCD) | Sensor:  Sentinel-2: Multispectral instrument (MSI).  Landsat-8: Operational Land Imager (OLI) |
| Coordinate system: ETRS89 LAEA | Data type:  Sentinel-2: TOA reflectances (Level 1), TOA radiances in sensor geometry (L1B)(Level 1) and BOA reflectances in cartographic geometry (L1C)(Level 2) |
| Production date: 22 - 03 - 2018 | Sensor resolution:  Sentinel-2: 10-60 m  Landsat-8: 30 m (visible, NIR & SWIR), 100 m (thermal) and 15 m (panchomatric) |
| Coverage (top L, BR coordinates): \* | Acquisition date: 2012 - 2015 |
| Grid size: 20 m | Grid size: - |
| Position accuracy: Less than one pixel | Positional accuracy: - |
| Vertical accuracy: - | Vertical accuracy: - |
| Completeness: Complete | |
| File type, format: Raster, TIFF image | |

Table 5: Technical specifications of Pan-European High Resolution Layers (HRL) dataset. Source: Deliverable D2.2

## Digital European elevation model

The Digital European elevation model (EU-DEM) is a digital surface model (DSM) of EEA member and cooperating countries representing the first surface as illuminated by the sensors. It is a hybrid product based on SRTM and ASTER GDEM data fused by a weighted averaging approach (European Environment Agency, Digital Elevation Model over Europe (EU-DEM), 2017). This is the v1.1 of EU-DEM, based on data acquired in 2011.

EU-DEM will be used as elevation source in order to identify mountainous and semi-mountanious marginal lands under MAIL’s scope.

| **Specification** | **Source data specification** |
| --- | --- |
| File name: Digital Elevation Model of Europe v1.1 | Sensor: (GLASS) Geoscience Laser Alimeter System |
| Coordinate system: ETRS89 LAEA | Data type: Level 1A, 1B, 2 and 3 data products |
| Production date: 20 - 04 - 2016 | Sensor resolution: 60-70 m x 60-70 m |
| Coverage (top L, BR coordinates): \*\* | Acquisition date: 2011 |
| Grid size: 25 m | Grid size: - |
| Position accuracy: - | Positional accuracy: - |
| Vertical accuracy: +/- 7 m RMSE | Vertical accuracy: - |
| Completeness: Complete | |
| File type, format: Raster, Geotiff 32 bits | |

Table 6: Technical specifications of Digital European elevation model dataset. Source: Deliverable D2.2

## Environmental dataset

In order to identify the legally protected areas, the European network of protected sites (Natura 2000) was used. Natura 2000 is the key instrument to protect biodiversity in the European Union. It is an ecological network of protected areas, set up to ensure the survival of Europe's most valuable species and habitats. Natura 2000 is based on the 1979 Birds Directive and the 1992 Habitats Directive. This version covers the reporting in 2018.

The European database on Natura 2000 sites consists of a compilation of the data submitted by Member States to the European Commission. This European database is generally updated once per year. (European Environment Agency, Natura 2000 data - the European network of protected sites, 2019)

| **Specification** | **Source data specification** |
| --- | --- |
| File name: Natura 2000 data - the European network of protected sites | Sensor: - |
| Coordinate system: ETRS89 LAEA | Data type: - |
| Production date: 12 - 04 - 2019 | Sensor resolution: - |
| Coverage (top L, BR coordinates): Europe | Acquisition date: 2018 |
| Grid size: 1 : 100.000 | Grid size: - |
| Position accuracy: - | Positional accuracy: - |
| Vertical accuracy: - | Vertical accuracy: - |
| Completeness: Complete | |
| File type, format: Vector, Shapefile | |

Table 7: Technical specifications of Environmental dataset. Source: Deliverable D2.2

## European Soil Database Derived data

In order to use soul variables, the European Soil Database Derived data compiled by the European Soil Data Centre (ESDAC) was used. On aforementioned database, a number of layers for soil properties have been created based on data from the European Soil Database in combination with data from the Harmonized World Soil Database (HWSD) and Soil-Terrain Database (SOTER). The available layers include:

* Total available water content
* Depth available to roots
* Clay content
* Silt content
* Sand content
* Organic carbon
* Bulk Density
* Coarse fragments

The layers of soil properties of Soil Typological Units (STUs) are only intended to facilitate modelling purposes. The final result of the modelling activity should be aggregated to SMUs or another larger mapping unit.

| **Specification** | **Source data specification** |
| --- | --- |
| File names: Various[[8]](#footnote-8) | Sensor: - |
| Coordinate system: ETRS89 LAEA | Data type: - |
| Production date: 2013 | Sensor resolution: - |
| Coverage (top L, BR coordinates): Europe | Acquisition date: 2013 |
| Grid size: 1 km | Grid size: - |
| Position accuracy: - | Positional accuracy: - |
| Vertical accuracy: - | Vertical accuracy: - |
| Completeness: Complete | |
| File type, format: Raster, Idrisi raster format | |

Table 8: Technical specifications of the European Soil Database Derived data dataset. Source: Deliverable D2.2

# Definition of GIS specifications

In the table below are summarized the original characteristic of each dataset used.

| **Dataset** | **Layer** | **File format** | **Projection** | **Coverage** | **Unit** |
| --- | --- | --- | --- | --- | --- |
| CORINE Land Cover | Land cover | Vector | ETRS89 LAEA | EEA 39[[9]](#footnote-9) | quantitative |
| Economic statistics | Gross domestic product | Table | - | Europe | € |
| NUTS3 | Vector | ETRS89 LAEA | Europe | No units |
| Pan-European High Resolution Layers | Tree Cover Density | Raster | ETRS89 LAEA | EEA 399 | % |
| Digital European elevation mode | EU-DEM | Raster | ETRS89 LAEA | EEA 399 | meters |
| Environmental dataset | Natura 2000 network | Vector | ETRS89 LAEA | Europe | quantitative |
| European Soil Database Derived data | 8 layers[[10]](#footnote-10) | Raster | ETRS89 LAEA | Europe | various |

Table 9: Main specifications of the datasets used. Source: Personal compilation

In order to homogenize and allow the required interoperability between each dataset, the following specifications should be followed.

## Coordinate Projection

For all layers the horizontal coordinate system shall be European Terrestrial Reference System 1989 (ETRS89) using Lambert Azimuthal Equal-Area projection (LAEA).

## File format

File format will be raster data (discrete) due to the fact that those are very useful for analysis and for storing data that varies continuously.

## Coverage

The extent used will be the 28 European Member States.

# Review of existing research projects and literature for available GIS models & workflows

As already presented on Deliverable 2.1, there is not a unique definition of marginal land. Therefore diverse methodologies for marginal land definition has been developed. Taking into account MAIL scope and the literature reviewed, the most suitable approach for marginal land identification is described as follows.

## Marginality index

In compliance with Bertaglia et al. (2007), a marginality index could be computed. The objective is to provide a quantitative operational tool, an integrated index on the basis of land cover data, as a measure of integrated marginality. This variable aims at synthesising information on the marginal characteristic of a region given its land-use pattern.

Relative marginality with respect to land use, is defined as those areas where there is relatively more non-productive and/or less productive land, compared both to land used for arable farming and to urban areas and transportation infrastructure (more productive classes)[[11]](#footnote-11). The method chosen is to compute a ratio of land-use type that can describe the relative abundance of marginal land-use in a certain area. As such, an area could be less marginal because it has a high proportion of high-intensity arable-land-based farming systems and/or a high proportion of urban or transport infrastructure. The ratio obtained is thus a good indicator of integrated relative marginality resulting from a different combination of factors.

Thus, the higher this ratio, the more the pattern of land use at the neighborhood area is considered as predominantly marginal. Computation of the marginality index will be performed on a moving window base, i.e. for each pixel of the land use raster, the aforementioned ratio was calculated taking into account the adjacent cells, in a square 3x3, 5x5 or 7x7[[12]](#footnote-12) (see figure below). In this way, the probability of a single pixel to be considered as marginal depends on the local land use pattern where the pixel is located.

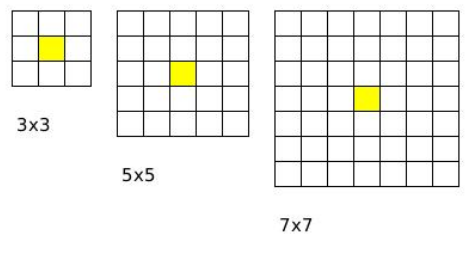


Figure 2: Different square areas where neighborhood analysis will be performed. In yellow is remarked the pixel that is computing the ratio. Source: personal compilation.

The implementation of the aforementioned index could be performed through a custom neighborhood analysis in QGIS (Tutić et al., 2018).

## Land use change

In order to include the dynamic aspect of marginality definition (see Figure 1) in the identification methodology, the change of land use will be included. During literature review two approaches in order to integrate dynamism were found:

1. CLC change layer. Under this approach CLC change layers of different periods could be used in order to calculate the probability of change of each use.
2. Change detection. Computation of Transition Potential matrix used for the prediction of land use change using CA-Markov model (Hamad et al., 2018; Dzieszko, 2014).

## Economic data

Gross domestic product (GDP) at current market prices by NUTS 3 regions could be used in order to define marginality index thresholds according to the regional economical status of each region using fuzzy logic (Van Echelpoel et al., 2015) or natural breaks (Chen at al., 2013).

## Tree Cover Density

Tree Cover Density (pan-European High Resolution Layer) will be utilized as forestry variable in order to select the marginal lands where activities such as forestation and reforestation will have more impact on the C stock accounting system as defined by the LULUCF directive[[13]](#footnote-13). According this regulation, not all of the forest related sinks will count toward the mitigation target (Grassi et al., 2019), therefore carbon stock contribution, for legal purposes, will depend if performed on already managed forest lands or afforested and forested lands. Distinction between those categories will be performed through TCD thresholds in order to select the most suitable marginal lands for MAIL project.

## Digital Elevation Model of Europe

The European Digital Model will be used for the selection or marginal lands located on mountainous or semi-mountainous areas. Thresholds for this variable should be set in a regionally taking into consideration the climate regions of Europe.

## Protected areas. European network of protected sites (Natura 2000)

This layer, where included all protected areas under Natura 2000 network, will be used as restriction layer in order to exclude from MAIL set of marginal lands those with any kind of protection status.

# Development of the final ML detection, criteria and classification scheme.

## ML definition refinement

In the previous work done in T2.1 “Literature review on Marginal Land definition”, in Deliverable 2.1, marginal lands for the MAIL project are defined as:

**Lands with significant, either environmental (biophysical variables) or socioeconomic, constraints and with potential to impact national accounting for C stock, excluding agricultural lands and other valuable areas (protected areas, uses with local importance etc.). Dynamic and variability are key concepts for marginal land identification.**

Examples of these areas include, but are not limited to, degraded and / or abandoned lands, lands with naturally low productivity due to biophysical constraints, and other degraded lands that have not (yet) been converted to other uses, e.g. post-industrial and post-mining sites.

This definition is further clarified in this document by taking into consideration the final indicators and criteria used in MAIL methodology for defining the MLs across Europe. Therefore, based on the indicators and criteria that will be analyzed in paragraphs 6.3 and 6.4, up until now, the definition can be modified to:

**Lands with significant, either environmental (biophysical variables) or socioeconomic, constraints and with potential to impact national accounting for C stock, excluding agricultural lands, *forest and impervious areas, permanent water and snow, peatbogs and marshes, as well as protected areas. MLs are further defined by constraints in a) terrain and soil physical variables such as slope, depth, texture, stoniness, drainage, water capacity, moisture, clay and sand, b) soil sustainability variables such as salinity, acidity, erosion, flood, sodicity, contamination, dryness and toxicity, c) soil productivity variables such as organic matter and cation exchange capacity.* Dynamic and variability are key concepts for marginal land identification.**

Nevertheless, for the definition of MLs in the MAIL project to be complete, there is still a need for defining the socioeconomic variables that will act as constraints.

## Methodology overview

Literature review done in T2.1 showed that most authors use a combination of land use/land cover and soil data to classify MLs. Additionally, climate, socio-economic and elevation data were regularly considered. The temporal dimension was investigated by evaluating multi-temporal datasets in order to assess land use and land productivity trends. Spatial classification was mostly implemented through geographic overlay of input data, using either a combination of binary constraints or fuzzy logic for ML classification. Constraints were divided into “soft” or “hard” constraints, “soft” constraints being factors with variable thresholds (e.g. elevation) and “hard” constraints binary exclusion factors, e.g. protected areas. Most factors indicative of marginality can be considered as “soft”. These include biophysical factors such as slope, elevation, soil quality/fertility and erodibility, which are inherent properties of the land or soil. Current land use and policy can be classified as “hard” constraints. Land that is currently in active use for agriculture cannot be seen as marginal, even if it has all characteristics of ML. This includes lands temporarily fallow as part of crop rotation. Protected areas are also excluded from marginal land classification within the MAIL project.

For the proper selection of indicators/criteria, along with the thresholds that will be used in the MAIL project, it was necessary to develop a methodological outline that will act as a roadmap for the next project tasks. The methodology that will be explained in the next pages (Figure 4) is still in an early stage and is likely to change due to potential obstacles in the implementation of it, as well as due to the addition of socio-economic and mountain indicators.

In general, mapping MLs in the MAIL project, will be performed in two phases. In the first phase, a top-down stepwise approach is followed (Figure 3), in which areas that are not MLs are incrementally removed, based on thresholds of various marginality criteria and indicators. Moving on, in phase two, the resulting MLs are classified into 3 classes, through a weighted overlay of the marginality indicators, similar to the approach followed by Li, Messina, Peter, & Snapp, (2017) and Zolekar & Bhagat, (2015).

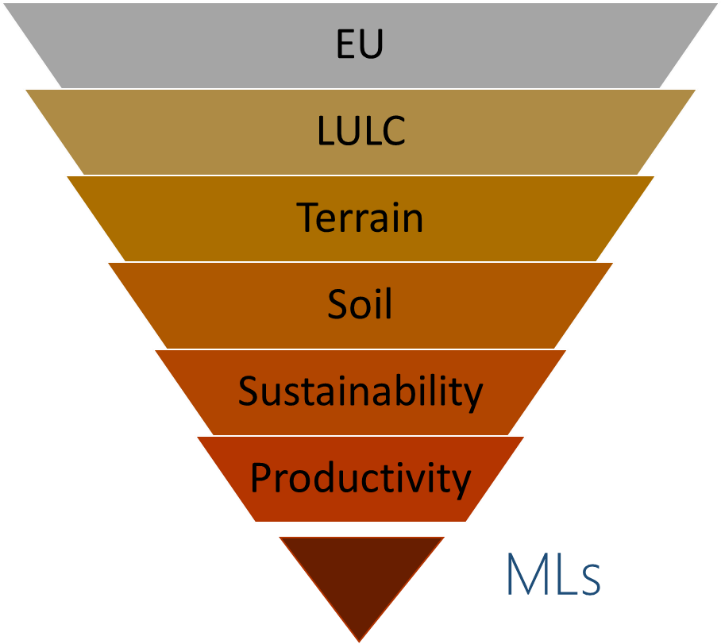


Figure 3: The **MAIL** methodology overview for mapping MLs.

In phase one, starting with the whole European area, areas are excluded, either based on land cover type (ex. urban areas, protected areas, water, forest areas, snow/ice etc.) (“hard” constraints), or based on thresholds of marginality indicators (ex. slope, moisture, salinity, productivity etc.) which were found in literature. The indicators/criteria used, are grouped by type, according to categorization made in D2.1, Table 8 (soil, climate, terrain, sustainability, productivity, LULC, socioeconomic).

In the second phase, each indicator that constitutes a “soft” constraint and has non-thematic data (numerical data), is ranked according to its importance (times found in literature) and weights are assigned to each one based on a Pairwise Comparison Matrix (PCM) of ranks (Zolekar & Bhagat, 2015). The remaining values of each indicator are grouped into 3 classes and are given scores, based on their contribution to marginality. This is followed by a weighted overlay in a GIS, based on the scores and weights of each indicator. Final step in the mapping of marginal lands, is the reclassification of the resulting product of the weighted overlay into 3 classes, depicting marginality: *1)* *Marginal lands with high plantation suitability, 2) Marginal lands with low plantation suitability* and *3) Potentially unsuitable lands*.

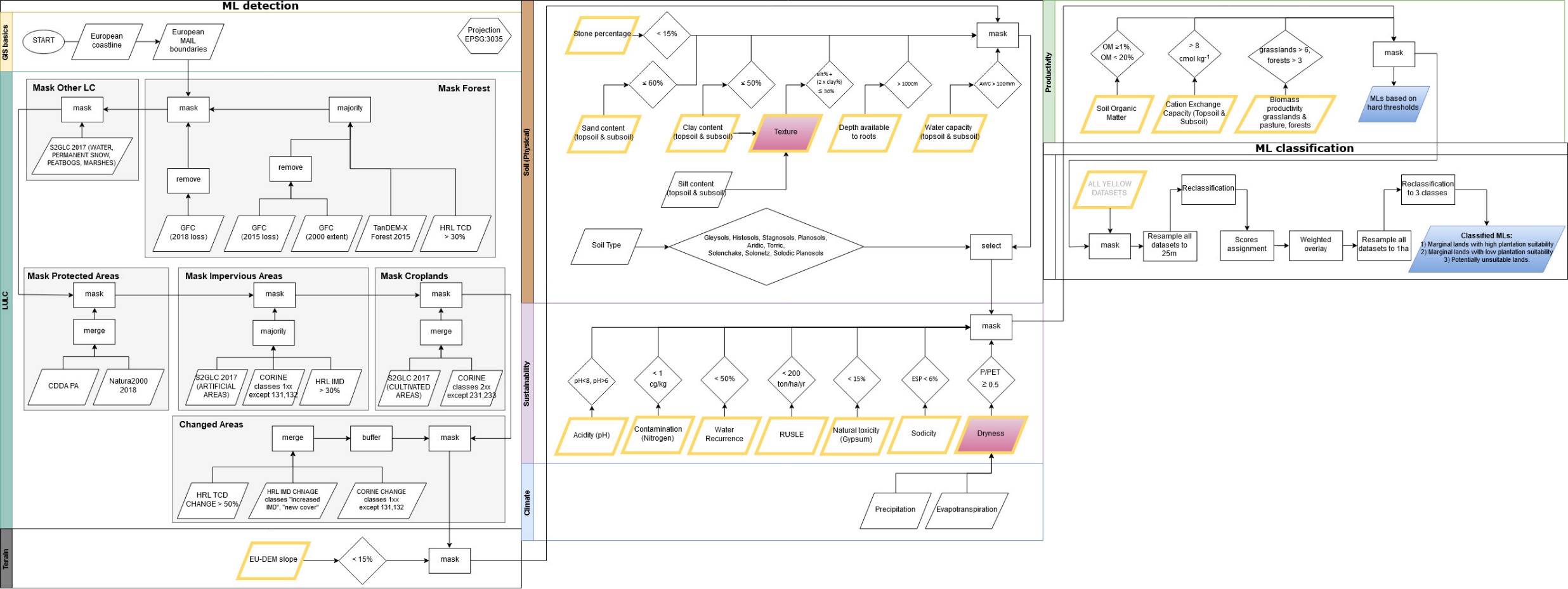


Figure 4: Workflow of the methodology for mapping and classifying MLs in MAIL. Yellow parallelograms depict datasets used in the weighted overlay, while pink parallelograms depict “synthetic” indicators/layers.

## Indicator / criteria selection

In general, the indicators and criteria for defining the MLs in the MAIL project, where selected by juxtaposing the indicators and criteria used in selected ML studies from T2.1 “Literature review on Marginal Land definition” and the available datasets from T2.2 “Collection of appropriate existing European/Global datasets”, taking into consideration the indicator units in both datasets and literature (Figure 5).

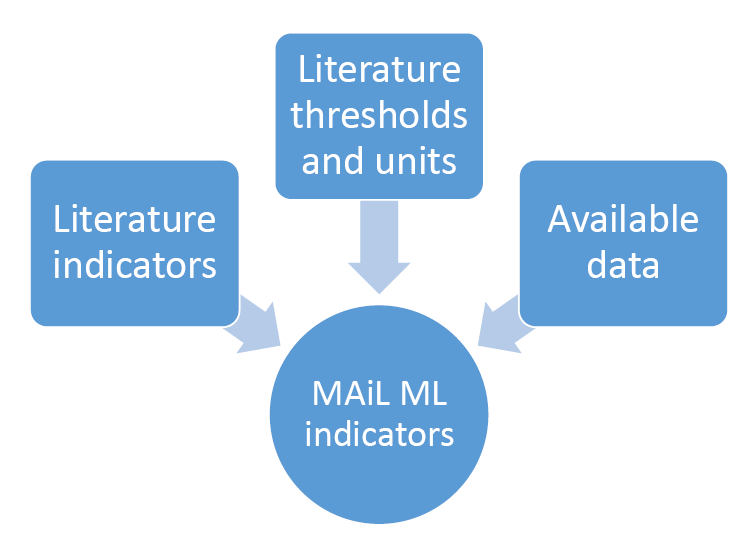


Figure 5: Overview of the indicator and criteria selection in the **MAIL** project.

A subset of 24 studies was made (Table 10), by selecting studies used in the literature review in T2.1 “Literature review on Marginal Land definition”. The subset contained studies only focusing on methodological aspects of mapping MLs. For each study, information regarding various methodological aspects where noted (study extent, MMU, technology used, datasets, creation of synthetic layers/indicators, ML classification scheme), as well as all the indicators used, grouped by type, according to categorization made in D2.1, Table 8 (soil, climate, terrain, sustainability, productivity, LULC, socioeconomic). Additionally, indicator thresholds were also noted, where a clear threshold was used.

Table 10: The subset of 24 studies on which the indicator selection was based on.

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Year** | **Authors** | **Study Title** |
| 1 | 2019 | Ciria et al. | Identification of Arable Marginal Lands under Rainfed Conditions for Bioenergy Purposes in Spain |
| 2 | 2018 | Peter et al. | Multiscalar approach to mapping marginal agricultural land: Smallholder agriculture in Malawi |
| 3 | 2018 | Vlachaki et al. | Final guidelines for the sustainable exploitation of Marginal Lands for bioenergy (D6.8) |
| 4 | 2018 | Gerwin et al. | Assessment and quantification of marginal lands for biomass production in Europe using soil-quality indicators |
| 5 | 2018 | Sallustio et al. | Assessing the economic marginality of agricultural lands in Italy to support land use planning |
| 6 | 2018 | Elbersen et al. | Mapping Marginal land potentially available for industrial crops in Europe |
| 7 | 2017 | Li et al. | Mapping Land Suitability for Agriculture in Malawi |
| 8 | 2016 | Ivanina et al. | Report on MagL concepts, debate and indicators (D2.3) |
| 9 | 2015 | Zolekar & Bhagat | Multi-criteria land suitability analysis for agriculture in hilly zone: Remote sensing and GIS approach |
| 10 | 2013 | Bandaru et al. | Soil Carbon Change and Net Energy Associated with Biofuel Production on Marginal Lands: A Regional Modeling Perspective |
| 11 | 2013 | Gelfand et al. | Sustainable bioenergy production from marginal lands in the US Midwest |
| 12 | 2013 | Kang et al. | Hierarchical marginal land assessment for land use planning |
| 13 | 2012 | Liu et al. | Bioenergy production potential on marginal land in Canada |
| 14 | 2011 | Cai et al. | Land Availability for Biofuel Production |
| 15 | 2011 | Gopalakrishnan et al. | A Novel Framework to Classify Marginal Land for Sustainable Biomass Feedstock Production |
| 16 | 2011 | Schweers et al. | Identification of potential areas for biomass production in China: Discussion of a recent approach and future challenges |
| 17 | 2010 | Eliasson et al. | Common criteria for the redefinition of Intermediate Less Favored Areas in the European Union |
| 18 | 2010 | James | Theory and identification of marginal land and factors determining land use change |
| 19 | 2009 | Milbrandt & Overend | Assessment of Biomass Resources from Marginal Lands in APEC Economies |
| 20 | 2008 | Bai et al. | Global Assessment of Land Degradation and Improvement 1. Identification by remote sensing |
| 21 | 2007 | Bertaglia et al. | Identifying European marginal areas in the context of local sheep and goat breeds conservation: A geographic information system approach |
| 22 | 2006 | Niu & Duiker | Carbon sequestration potential by afforestation of marginal agricultural land in the Midwestern U.S. |
| 23 | 2005 | Roehrig et al. | The Determination of Natural Agricultural Potential in Western Africa Using the Fuzzy Logic Based Marginality Index |
| 24 |  | Elbersen et al. | Deliverable 2.6 Methodological approaches to identify and map marginal land suitable for industrial crops in Europe |

At the current stage of the project, emphasis was given to soil, climate, terrain, sustainability, productivity and LULC constraints, leaving socioeconomic aspects and the mountainous/semi-mountainous variable, which are more complex criteria, aside for the moment. In the following paragraphs, indicators and thresholds are reviewed. Table 11, Table 12 and Table 13 summarize the final indicators chosen for soil, climate, terrain, sustainability and productivity constraints which are analyzed in paragraph 6.4.

Table 11: Soil and terrain marginality indicators used in *MAIL*.

|  |  |
| --- | --- |
| **Indicator** | **Times Found in Literature** |
| slope | 18 |
| depth available to roots | 18 |
| texture | 9 |
| stoniness | 8 |
| drainage | 7 |
| water capacity | 6 |
| moisture | 5 |
| clay | 4 |
| sand | 4 |

Table 12: Sustainability marginality indicators used in *MAIL*.

|  |  |
| --- | --- |
| **Indicator** | **Times Found in Literature** |
| salinity | 10 |
| acidity (pH) | 9 |
| erosion | 8 |
| flood | 6 |
| sodicity | 5 |
| contamination | 4 |
| dryness | 2 |
| natural toxicity | 1 |

Table 13: Productivity marginality indicators used in *MAIL*.

|  |  |
| --- | --- |
| **Indicator** | **Times Found in Literature** |
| soil organic matter | 8 |
| cation exchange capacity | 4 |
| productivity | 3 |

## Initial indicator classification categories and thresholds

In general, the indicators and criteria for defining the MLs in the MAIL project, where selected by combining the indicators and criteria resulting from the literature review with the available datasets from T2.2, taking into consideration the indicator units in both datasets and literature (Figure 5).

In all indicator cases, when more than one threshold was observed, the final threshold chosen for the MAIL project was the one producing larger MLs area. Following this, a match was tried to be achieved between each indicator to a corresponding dataset from T2.2. In cases where a match could not be found between an indicator and a dataset (conceptually and in units), the indicator was omitted. For example, no dataset existed for “Phosphorus level”. Evidently, the study in which this indicator was used (Zolekar & Bhagat, 2015), used national data for the country studied which is not available in a European level.



### Land Use/Land Cover indicators

Many authors have used Land Use/Land Cover (LULC) in detecting MLs (Bai, Dent, Olsson, & Schaepman, 2008; Bertaglia, Joost, & Roosen, 2007; Gelfand et al., 2013; Gerwin et al., 2018; Niu & Duiker, 2006). LULC is considered as a single indicator with multiple sub-variables that form the “hard” constraints to marginality. LULC datasets are thematic datasets and thus thresholding is in the form of masking.

In the MAIL project, LULC datasets are used in phase one. Starting with the whole European area, areas are excluded, initially based on land cover type. More specifically, areas excluded from the analysis, based on LULC are: forest areas, croplands, impervious areas, protected areas, water, permanent snow, marshes and peatbogs.

Moreover, to incorporate the dynamic aspect of MLs, changed areas are also considered. In particular, Tree Cover Density (TCD) will be utilized as forestry variable in order to select the marginal lands where activities such as aforestation and reforestation will have more impact on the C stock accounting system. In addition, increase in imperviousness over the past years will also be considered, by masking around a buffered zone of areas marked as such.

### Terrain and Soil indicators

#### Slope:

Slope is the angle the soil surface makes with the horizontal, expressed in degrees or as a percentage. It is considered as one of the main criteria for MLs. Not only it affects water drainage and erosion, but it is also associated with shallow soils and mechanical constraints in agriculture (Eliasson et al., 2010). Sallustio et al. (2018), used a >30% slope threshold for unsuitable agricultural lands in Italy. Elbersen et al. used a threshold of >17.5 degr (30%) in pair wise combinations for mapping MLs suitable for growing industrial crops. Marginally suitable agricultural lands had slopes 6-16% and marginally unsuitable 16-30% in Li, Messina, Peter, & Snapp, (2017). Ivanina, Roik, & Hanzhenko, (2016) excluded areas with slopes ≥ 15% for mapping agricultural ML suitable for bioenergy crops. Zolekar & Bhagat, (2015) considered slopes 12-20 degr (21-36%) marginally suitable for agriculture, while slopes >20 degr (36%), unsuitable. For Gelfand et al., (2013), MLs have slopes <20%. Marginal agricultural land for biomass feedstock production was mapped, by selecting areas with slope >15% in Gopalakrishnan, Cristina Negri, & Snyder, (2011). Schweers et al., (2011) marked slopes >25 degr (47%), as unsuitable for agriculture. Unfavorable agriculture areas in the EU have slopes >15%, according to Eliasson et al., (2010). Finally, lands with moderate (8-16%) or steep slope (16-30%), are considered as marginal for biomass production in Milbrandt & Overend, (2009).

In the MAIL project the threshold chosen was the smallest one, in other words MLs in MAIL have slopes ≥ 15%.

#### Rooting depth:

Rooting depth is the depth from the soil surface to a hard rock, in which plant roots can grow. Apart from rooting depth being a constraint itself, shallow soil it affects also water capacity (Zolekar, 2018) and is thus a major factor for characterizing a land as marginal. Ciria, Sanz, Carrasco, & Ciria, (2019) used a threshold of 100cm to map arable MLs bellow this. Elbersen et al. used depths <30cm in pair wise combinations for mapping MLs suitable for growing industrial crops. Li et al., (2017) classify soil depth in 5 classes, with MLs having depths <50cm. Ivanina, Roik, & Hanzhenko, (2016) excluded areas with rooting depth <35cm for mapping agricultural ML suitable for bioenergy crops. Marginally suitable agriculture soils were the ones with depths 30-50cm in Zolekar & Bhagat, (2015). In a study by Kang et al., (2013), MLs had soil depths < 50cm while in Eliasson et al., (2010), unfavorable agriculture areas in the EU have rooting depths <30cm. Milbrandt & Overend, (2009) used depths of <50cm to map marginal lands for biomass production.

In the MAIL project the threshold chosen was the largest one, in other words MLs in MAIL have rooting depths <100cm.

#### Soil texture (soil structure/substrate):

Soil texture (also found as “soil structure” or “substrate”), refers to the relative proportions of different-sized soil particles, consisting of sand, silt, and clay. It controls the soil structure, fertility, and water availability (Eliasson et al., 2010; Zolekar, 2018) and is an important soil marginality indicator. Soils with high sand, clay or rock proportions are unfavorable for growing crops (Ivanina et al., 2016). For this indicator many authors use thresholds in percentages of soil particles while others use soil type as a proxy for texture. Ciria et al., (2019) use stoniness >15% to characterize lands as marginal. Elbersen et al., used a combination threshold of silt and clay (silt% + (2 x clay%) ≤30%), sand <60% and stoniness>35%, for ML mapping. Eliasson et al., (2010) used percentages of soil particles (>60% clay and >15%(v/v) of coarse fragments), but also soil type classes (clay, silty clay, or sandy clay with vertic properties) to detect MLs. Soil type was used as a proxy for texture in Li et al., (2017) to detect agriculture MLs, by selecting Sandy-Clay and Clay soil types in Malawi. Zolekar & Bhagat, (2015), used Loam Soil texture class for marginally suitable agriculture areas and last, Milbrandt & Overend, (2009), used coarse textured or sandy soils (Arenosols, Regosols, and Vitric Andosols with coarse texture, and all soils with petric and stony phase), to map marginal lands for biomass production.

In the MAIL project the percentage thresholds of different soil particles were chosen. Texture is considered as a “synthetic” indicator which depends on several other indicators (stoniness, silt, clay, sand). Thus, MLs in MAIL have stoniness% >15%, silt% + (2 x clay%) ≥ 30%, clay ≥50% and sand >60%.

#### Soil drainage:

Soil drainage refers to the air (and thus oxygen) supply in soil pores by removal (or non- addition) of water (Eliasson et al., 2010; Li et al., 2017). It is usually assessed by soil type. Marginal lands in Elbersen et al. have Gleysols, Histosols, Stagnosols, Planosols, and soils with Histic, Gleyic and Stagnic primary qualifiers. Gleysols and Stagnosols are also found in Eliasson et al., (2010) for unfavorable agriculture areas in the EU. In some studies, drainage is found in local/state data as a variable per se (ex. MASDAP soil data record in Li et al., (2017), USDA-NRCS in Kang et al., (2013).

The only available datasets found in D2.2, associated with soil drainage throughout whole Europe, were soil type datasets. Therefore, in the MAIL project soil types Gleysols, Histosols, Stagnosols, and Planosols were chosen as indicator variables for MLs.

#### Water storage capacity (maximum water holding capacity/profile available water):

Water storage capacity of soils determines soil depth, leaching process of nutrients and pesticides, the water availability of a soil profile available for vegetation, the cropping pattern, irrigation facilities, etc, and is a crucial parameter of soil quality, in particular in regions of rain deficit during the vegetation period (Mueller, Schindler, Behrendt, Eulenstein, & Dannowski, 2007; Zolekar & Bhagat, 2015). MLs in Ivanina et al. (2016) had a level of underground water in 0-80 cm, and a gleyic pattern in 0-40 cm, according to field surveys. Marginally suitable agriculture lands had Maximum Water Holding Capacity MWHC 100-200 and unsuitable lands MWHC <100 in Zolekar & Bhagat, (2015). Kang et al. (2013) used the water table data by FAO as a marginality indicator with MLs having a water table <30cm.

As none of the variables used in literature could match a European dataset from the MAIL task 2.3, it was decided to use the Available Water Capacity (AWC) from the ESDAC dataset “Topsoil physical properties for Europe (based on LUCAS topsoil data)”. AWC in this dataset was derived as the difference between the − 33 kPa and the − 1500 kPa water content (expressed as volume fraction). The threshold chosen was AWC <100mm, based on Zolekar & Bhagat, (2015).

#### Soil moisture:

Soil moisture is measured as the water availability during the plant’s growing period, which is crucial for plant normal growth and development and is thus an important ML indicator (Eliasson et al., 2010). Soil moisture is linked with numerous sub-factors such as length of rowing period, temperature, precipitation, evapotranspiration, soil water content and finally soil type. Eliasson et al. (2010), analyzed the number of days within the growing period (as defined by temperature >5°C), for which the amount of precipitation and water available in the soil profile exceeds half of potential evapotranspiration. Areas where the number of days is less than 90 are considered as unfavorable. Marginal lands had water content in the soil exceeding field capacity for at least 210 days in Elbersen et al. On the other hand, Zolekar & Bhagat (2015), classified NDWI to four classes (good, medium, less, very less), with class “less” being marginally suitable for agriculture. Last, in a more simple approach, Kang et al. (2013) defined soil moisture by soil type, with dry soils Aridic and Torric being an indicator for MLs.

In the MAIL project MLs have low soil moisture and are defined by soil types Aridic and Torric.

#### Soil type:

Soil type is an indicator widely used as a marginality indicator, although it is mainly connected to other factors. Soil type is associated to *drainage* (Gleysols, Histosols, Stagnosols, Planosol, Soils with primary qualifiers Histic, Gleyic and Stagnic), in Elbersen et al. In Milbrandt & Overend, (2009), marginal land soil types are associated to numerous factors: *texture* (Arenosols, Regosols, and Vitric Andosols with coarse texture, and all soils with petric and stony phase), *salinity* (Solonchaks, Solonetz, and Solodic Planosols), *gypsum levels* (Gypsic Xerosols and Gypsic Yermosols), *calcium levels* (Calcisols) etc. In Kang et al., (2013), soil type is used as a proxy for *moisture* in MLs (Aridic, Torric). Soil type is used as an indicator per se in another study, for mapping agricultural ML in Spain (Entisol and Aridisol soil types) (Ciria et al., 2019).

Because soil type is used for defining other indicators in the MAIL project (drainage, moisture, salinity), it was not used as an indicator per se.

Table 14: Terrain and soil indicators and thresholds, sorted by importance (times used in literature review)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Indicator** | **Data source** | **MAIL threshold** | **MAIL threshold based on** | **Literature Thresholds** |
| slope | EU-DEM 2016 25m | ≥ 15% | (Gopalakrishnan et al., 2011) | > 15% - 47% |
| depth available to roots | ESDAC 2013 1km | < 100cm | (Ciria et al., 2019) | < 30cm – 100cm |
| texture | ESDAC 2013 1km | silt% + (2 x clay%) ≥ 30% | (Elbersen et al., n.d.) | silt% + (2 x clay%) ≥ 30% |
| stoniness | ESDAC 2008 | > 15% | (Ciria et al., 2019) | > 15% |
| drainage | WRB “Soil type” 2008 | By soil type: Gleysols, Histosols, Stagnosols, Planosols | (Eliasson et al., 2010) | By soil type: Gleysols, Histosols, Stagnosols, Planosols |
| water capacity | ESDAC “Total available water content from PTF” 1km 2013 | AWC < 100mm | (Zolekar & Bhagat, 2015) | AWC < 100, water table < 0.3m |
| moisture | WRB “Soil type” 2008 | By soil type: Aridic, Torric | (Kang et al., 2013) | Various calculation methods and thresholds. |
| clay | ESDAC 2013 1km | ≥ 50% | (Eliasson et al., 2010) | ≥ 50%, >60% |
| sand | ESDAC 2013 1km | > 60% | (Elbersen et al., n.d.) | > 60% |

### Sustainability indicators

#### Soil salinity/alkalinity:

Soil salinity is the presence of salts of alkalis (sodium, potassium, magnesium and calcium) on the land surface, in soil. Soil salinity can be caused by environmental factors or human induced factors that disturb natural ecosystems. Soil salinity affects plant productivity and soil structure, leading to the creation of toxic substances and serious soil erosion (Eliasson et al., 2010), and is thus a major marginality indicator. Salinity levels above 6 dSm-1, make winter cereal growth impossible for Spain in the study by (Ciria et al., 2019), whereas in Elbersen et al., Solonchaks and soils with a salic qualifier, soils with salt levels >15 dSm-1 in more than 50% of the mapping unit area were used in pair wise combinations for mapping MLs suitable for growing industrial crops. Soil type is again used for characterizing salinity, with saline soils being Solonchaks, Solonetz, and Solodic Planosols (Milbrandt & Overend, 2009).

Due to lack of European datasets associated with salinity in the units described in literature, in the MAIL project, salinity is described by Solonchaks, Solonetz, and Solodic Planosols soil types.

#### Soil acidity (potential of hydrogen/pH):

Acidification is a major soil degradation factor and affects soil fertility. Acidity is measured by determining the pH of a soil (Mueller et al., 2007), and is a major marginality indicator. In Kang et al. (2013), marginal areas are characterized by pH levels in soil > 9 or <5.5. Elbersen et al. used values of pH in topsoil water <5, for mapping MLs suitable for growing industrial crops. Similarly, Ivanina et al. (2016) used values of pH in topsoil water ≤5.5 for mapping agricultural ML suitable for bioenergy crops. Marginal lands for biomass production had pH values 5.5 - 4.5 or <4.5 in (Milbrandt & Overend, 2009), and finally, Ciria et al. (2019) used pH values <6 and >8 to map arable agriculture MLs.

In the MAIL project, soft pH thresholds were chosen, in order to result in larger areas as potential MLs, i.e. pH >8 and pH <6.

#### Soil erosion:

Being a major hazard, soil erosion is inseparably connected to slope and soil texture, but is also affected by climate, vegetation cover and agricultural practices (Mueller et al., 2007). Li et al. (2017) used erosion as a marginality indicator by calculating the Universal Soil Loss Erosion and the revised version (RUSLE), classified in 11 classes. Marginally suitable agricultural lands had moderate erosion risk (aprox. 200-1500 ton/yr/ha), and marginally unsuitable agricultural lands had higher erosion risk (aprox. 1500-2500 ton/yr/ha). Zolekar & Bhagat, (2015) used elevation, classified in 3 classes, as a proxy to depict soil erosion risk, while Roehrig & Menz (2005), in this same context, used slope. Both soil erosion tolerance rate of 2 ton/ha, and erosion K factor K=8 were used to assess marginal lands in (Kang et al., 2013). Lastly, Elbersen et al. used the classes High’ and ‘Very High’ of the Index of Land Susceptibility to Wind Erosion (ILSWE), as well as the WaTEM/SEDEM spatial database for sensitivity for erosion by water (> 100 ton/ha/yr) as thresholds for MLs suitable for growing industrial crops.

The MAIL threshold for soil erosion is based on the units in the European RUSLE available dataset and the threshold value from (Li et al., 2017), i.e. RUSLE > 200 ton/ha/yr.

#### Flooding (ponding):

Flooding is considered as a marginality index by some authors, since flooding implies a risk to most plant species. Elbersen et al. used a threshold of >1-2m flood in a 2 year return time during the growing season, for mapping MLs suitable for growing industrial crops. Flood duration and frequent ponding are also used by Kang et al. (2013) for assessing MLs. Gopalakrishnan et al., (2011) classified lands as marginal for biomass production when chance of flooding in a year exceeded 50%.

Based on the last study, in the MAIL project, the threshold for describing a land as marginal, decided to be a recurrence of water >50% from JRC Global Surface Water. Recurrence provides information concerning the inter-annual behavior of water surfaces and is defined as the frequency with which water returns from years to year, expressed as a percentage (JRC, 2018).

#### Sodicity:

Soil sodicity is a characteristic of land for which the proportion of adsorbed sodium in the soil clay fraction is too high for plants to perform or survive. The effects of sodicity are often indirect as they affect vital soil properties rather than plant growth itself (Eliasson et al., 2010). Kang et al. (2013), classified lands as marginal when Na saturation exceeded 6–15%. Elbersen et al. and Eliasson et al. (2010) used Saturation with Exchangeable Sodium (ESP) to assess sodicity (ESP >15% in more than 50% of the mapping unit area and ESP >6, respectively).

In MAIL, ESP >6 was chosen as a threshold for sodicity in MLs.

#### Toxicity from contamination:

Soil toxicity by pollutants other chemicals is a serious obstacle for conventional agriculture because of increasing risks of phytotoxicity and food contamination. Areas with high toxicity from contamination are usually referred to as marginal (Ivanina et al., 2016). One of the most widespread contaminants is Nitrate, with Nitrate ≥10mg L−1 in groundwater being the lower limit of classifying a land as marginal in some studies (Gopalakrishnan et al., 2011; Ivanina et al., 2016). This threshold is also used in MAIL, by slightly changing the units to match the corresponding dataset (1cg/kg).

#### Dryness (Aridity Index):

Dryness is defined as an overall low soil water content, resulting from a natural imbalance in the water availability (through low annual precipitation and high annual evaporative demand) (Ivanina et al., 2016). Elbersen et al. and Ivanina et al. (2016) used dryness as a marginality indicator, by computing the Aridity Index (AI). The AI is the ratio of the total annual precipitation (P) to the total annual potential evapotranspiration (PET), and it expresses the relationship between severity of dry conditions and biomass production. Severe conditions correspond to AI values ≤0.5. In (Bai et al., 2008), for calculating the AI, PET is defined as PET = P/ (0.9 + (P/ L)2 ), and L = 300 + 25T + 0.05T3, where T is mean annual temperature and P is annual precipitation in mm. AI is also used by Cai, Zhang, & Wang (2011), but with a different formula.

In MAIL, the AI was chosen for assessing the dryness indicator with MLs having AI ≤0.5.

#### Natural toxicity:

Natural toxicity is an adverse chemical condition for soils, affecting soil fertility, caused by chemical elements that naturally exist such as Aliminium, Sulfur or Gypsum. Eliasson et al. (2010) marks lands as unfavorable when Gypsum presence >15%. This threshold is also used in MAIL.

Table 15 Sustainability indicators and thresholds, sorted by importance (times used in literature review)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Indicator** | **Data source** | **MAIL threshold** | **MAIL threshold based on** | **Literature Thresholds** |
| salinity | WRB “Soil type” 2008 | By soil type:  Solonchaks, Solonetz, and Solodic Planosols | (Milbrandt & Overend, 2009) | >5 dSm-1, >6 dSm-1, Soil type (Solonchaks, Solonetz, and Solodic Planosols) |
| acidity (pH) | ESDAC 2009 | pH>8, pH<6 | (Ciria et al., 2019) | pH>9 or <5.5, pH<5, 5.5 - 4.5 or <4.5, pH<6 and pH>8, ≤ 5.5 |
| erosion | RUSLE 2015 100m, ESDAC 2015 | RUSLE > 200 ton/ha/yr | (Li et al., 2017) | By elevation, by slope, Soil erosion tolerance rate 2 ton/ha, Erosion by water > 100 ton/ha/yr |
| flood | JRC Global Surface Water- “recurrence” 1984-2018 30m | > 50% | (Gopalakrishnan et al., 2011) | >1-2 m flood in 2yr, >50% chance of flooding per year |
| sodicity | WISE 2015 10km | > 6% | (Eliasson et al., 2010) | >6–15% , > 15% ESP in more than 50% of the mapping unit area |
| contamination | WISE 2015 10km | Nitrogen > 1cg/kg | (Gopalakrishnan et al., 2011; Ivanina et al., 2016) | nitrate ≥ 10 mg L−1 |
| dryness | TerraClimate “Precipitation accumulation” and “evapotranspiration” 2018 4.5km | P/PET ≤ 0.5 | (Elbersen et al.; Ivanina et al., 2016) | P/PET ≤ 0.5 |
| natural toxicity | WISE 2015 10km | gypsum > 15% (for WISE GYPS > 150 g/Kg) | (Eliasson et al., 2010) | gypsum > 15% |

### Productivity indicators

#### Soil Organic Matter (organic carbon):

According to the United States Department of Agriculture (USDA), “*Soil organic matter (SOM) is the organic component of soil, consisting of three primary parts including small (fresh) plant residues and small living soil organisms, decomposing (active) organic matter, and stable organic matter (humus).*” SOM is a reservoir of nutrients for plants, inextricably related to soil fertility, and is thus used in many studies as a marginality indicator. Although some studies use only low SOM values to depict MLs (<1% SOM in the study of Ciria et al., (2019) and <0.8% SOM in Li et al., (2017)), others also used the high values of SOM to take histic soils into consideration. Elbersen et al. used a threshold of SOM <0.75% (at depth range 0-30 cm) for low fertility MLs and SOM ≥20% to express Histosols (Peat soils). Eliasson et al. (2010) used high values (SOM >30%) to assess histic soil areas as less favorable in the EU.

In the MAIL project, soft SOM thresholds were chosen, in order to result in larger areas as potential MLs, i.e. SOM <1% and SOM ≥ 20%.

#### Cation Exchange Capacity (CEC):

According to Li et al. (2017) “*Cation exchange capacity (CEC) is the negatively charged content (usually clay and organic matter particles) in soils that can hold positively charged ions such as calcium, magnesium and potassium through electrostatic forces. It is used as a measure of fertility and nutrient retention capacity”*. In the same study, it was found that MLs had low CEC, namely CEC <8cmol kg-1). In (Kang et al., 2013) though, a lower threshold was used of CEC <4 meq./100g at pH=7.

For MAIL, like in all cases, the threshold that results in larger areas as potential MLs was chosen, i.e. CEC <8 cmol kg-1.

#### Productivity:

*Crop* productivity, as an indicator for defining ML, is used mostly by creating a threshold for a main crop type cultivated in the study area. Ciria et al. (2019), found that annual productivity of ≤1.5Mgha-1 of wheat and barley was insufficient to cover the calculated cultivation costs, and thus land with productivity bellow this is considered as marginal. Bandaru et al. (2013), used the National Commodity Crop Productivity Index to classify lands as marginal for NCCPI 0.42–0.14. Lands with productivity for the main grain crop (nonirrigated yields of corn) <9 tonnes/ha (4 tons/acre), are defined as marginal in Gopalakrishnan et al. (2011). On the other hand, in (Cai et al., 2011), MLs are classified according to *land* productivity, and productivity is used as a synthetic indicator according to factors such as slope, soil type, soil temperature regime etc, based on land cover types.

It is evident that, crop productivity is tightly related to crop type and on national data, and thus can’t be used in the MAIL project due to unavailable data. Nevertheless, based on the ESDAC dataset “Soil Biomass Productivity maps of grasslands and pasture, of croplands and of forest areas in the European Union”, productivity as an indicator will be used for grasslands, pastures and forest areas, taking as a threshold the lowest value classes.

#### Net primary productivity (NPP):

Net primary production, or Net primary productivity (NPP), is the difference between total photosynthesis and total plant respiration in an ecosystem (Clark et al., 2001). Peter, Messina, & Snapp, (2018) used interannual productivity (from MODIS NPP data), for classifying agricultural areas as marginal or not. Roehrig & Menz (2005) used NPP in a fuzzy logic marginality index. Furthermore, NPP is used for measuring land degradation, in relation to other factors in the study by Bai et al. (2008).

Due to no clear threshold used in the examined literature, it was decided not to use NPP in the MAIL project as a marginality indicator.

#### Soil fertility:

Soil fertility, being a complex variable, dependent on many other factors and other marginality indicators, is also used as an independent indicator in some studies. Soil fertility was assessed using the Soil Quality Rating (SQR) System in (Mueller et al., 2007) and (Ivanina et al., 2016), by taking field soil samples. More specifically, in the last, regions with SQR ≤ 25 are considered as marginal. Fertility factor Sf was chosen from the Leemans and van den Born database of soil properties in (Roehrig & Menz, 2005). Finally, soil fertility is used as a combination of soil pH and SOC content for defining MLs in Elbersen et al.

In three out of four studies reported above, either there is no clear threshold, either local datasets are used. In the study by Elbersen et al., the variables used (pH, SOC) will be used in MAIL as independent indicators. Therefore, it was decided to omit the soil fertility indicator from the analysis.

Table 16 Productivity indicators and thresholds, sorted by importance (times used in literature review)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Indicator** | **Data source** | **MAIL threshold** | **MAIL threshold based on** | **Literature Thresholds** |
| soil organic matter | LUCAS “Topsoil Soil Organic Carbon” 2015 1km | OM <1%, OM ≥ 20% | (Ciria et al., 2019; Elbersen et al.) | OM <1%, OM < 0.75%, OM ≥ 20%, OM ≥ 30%, Histosols |
| cation exchange capacity | WISE 2015 10km | CEC < 8 cmol kg-1 | (Li et al., 2017) | CEC<8 cmol kg-1, CEC<4 meq/100g at pH=7 |
| productivity | “Soil biomass productivity of grasslands and pastures” “Soil biomass productivity of forest areas” 2016 1km | grasslands < 6, forests < 3 | - | No clear threshold. |

### Climate indicators

#### Precipitation:

Mostly used for calculating other indicators such as dryness and aridity, precipitation is used in many studies as a marginality indicator per se.Rain fed arable land with annual precipitation less than 400mm is marginal according to Ciria et al. (2019). Peter et al. (2018) used accumulated growing season precipitation <750mm and >1217mm to define areas as marginal for maize growth. In (Li et al., 2017), annual precipitation data for Malawi were obtained by accumulating mean monthly rainfall data. Average annual precipitation is used, not as a marginality indicator, but as a suitability indicator for bioenergy crops (300-1000mm) in the study of Ivanina et al. (2016). In another approach, internal variability of seasonal precipitation, based on negative anomalies of monthly growing season precipitation was used, along with a fuzzy logic marginality indicating system, in (Roehrig & Menz, 2005). Finally, similar to the method followed in MAIL, precipitation is used for calculating the “dryness” marginality indicator Elbersen et al., and to calculate the Global Aridity Index (Bai et al., 2008; Cai et al., 2011), as already described in 6.4.3.

In MAIL, precipitation is used only as a variable for calculating the synthetic dryness/aridity indicator.

#### Evapotranspiration:

The United States Geological Survey (USGS) defines evapotranspiration as “*the sum of evaporation from the land surface plus transpiration from plants*”. In all the reviewed studies of this document, evapotranspiration is used for calculating other marginality indicators such as dryness/aridity (Bai et al., 2008; Cai et al., 2011; Elbersen et al.; Ivanina et al., 2016).Based on these, the same approach is used in in the MAIL project, thus evapotranspiration was used for calculating the dryness/aridity indicator and applying the threshold accordingly.

#### Air temperature:

Like precipitation, air temperature is also used for calculating other indicators such as length of growing period, aridity index, energy-use efficiency etc. In Elbersen et al., temperature is used to calculate the Length of Growing Period (LGP), as the days with average temperature >5. If LGP ≤180 days, then the area is considered as ML. In (Cai et al., 2011), as mentioned above, mean monthly temperature is used to calculate the global aridity index. Moreover, it was found that air temperature correlates highly with land surface temperature (LST) so in this study, air temperature was used instead of LST for the fuzzy logic modeling for ML detection. In (Bai et al., 2008), annual accumulated temperature is used to calculate energy-use efficiency (EUE) (ratio of annual sum NDVI to annual accumulated temperature). Areas with a negative EUE indicator are marked as degraded.

Due to no clear threshold used in the examined literature, it was decided not to use air temperature in the MAIL project as a marginality indicator.

Table 17 Climate indicators and thresholds, sorted by importance (times used in literature review)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Indicator** | **Data source** | **MAIL threshold** | **MAIL threshold based on** | **Literature Thresholds** |
| precipitation | TerraClimate “precipitation accumulation” 2018 4.5km | Calculation of other indicators | - | Pannual,sum<400mm, PGS,sum<750mm and PGS,sum>1217mm |
| evapotranspiration | TerraClimate “evapotranspiration” 2018 4.5km | Calculation of other indicators | - | Calculation of other indicators. |

## ML classification schemes

The term marginal land does not have a unique definition. Categorization of MLs in literature is based on the definition of MLs used by each approach. Therefore, there are different classifications depending on each study’s goals. The MAIL project aims to map MLs which can potentially become carbon sinks, thus MLs must be classified in Carbon sequestration capacity groups based on the indicators. Therefore, MLs are reclassified into 3 classes, depicting marginality: *1)* *Marginal lands with high plantation suitability, 2) Marginal lands with low plantation suitability* and *3) Potentially unsuitable lands*.

The indicators that used to derive the final MLs classes are the non-thematic data ones (numerical data indicators). A rank is given to each indicator (Table 18), according to its importance (times found in literature). Alternatively, the rank given to each indicator can be based on expert opinions in Carbon sequestration. Subsequently, the Pairwise Comparison Matrix (PCM) of ranks (Zolekar & Bhagat, 2015) is calculated (Table 19), along with the normalized PCM, and the weight for each indicator. The cell values of PCM were divided by sum of the column to obtain the cell values in the normalized PCM and averaged in row to calculate the weights of each indicator (Table 20). The calculated weights are scaled from 0 to 1 in ascending order, and have a sum equal to 1, to maintain hierarchy according to their importance in marginality. The remaining values of each indicator are grouped into 3 classes and are given scores, based on their contribution to marginality (Table 21, Table 22, Table 23). This is followed by a weighted overlay in a GIS, based on the scores and weights of each indicator. Final step in the mapping of marginal lands, is the reclassification of the resulting product of the weighted overlay into the 3 aforementioned ML classes.

Table 18: The ranks given to each indicator, in order to calculate the weights.

|  |  |  |
| --- | --- | --- |
| **Indicator** | **Times Found in Literature** | **Rank** |
| slope | 18 | 1 |
| depth available to roots | 18 | 1 |
| acidity (pH) | 9 | 2 |
| texture | 9 | 2 |
| erosion | 8 | 3 |
| stoniness | 8 | 3 |
| soil organic matter | 8 | 3 |
| water capacity | 6 | 4 |
| flood | 6 | 4 |
| sodicity | 5 | 5 |
| clay | 4 | 6 |
| sand | 4 | 6 |
| contamination | 4 | 6 |
| cation exchange capacity | 4 | 6 |
| productivity | 3 | 7 |
| dryness | 2 | 8 |
| natural toxicity | 1 | 9 |

Table 19: The pairwise comparison matrix of the marginality indicators, based on the ranks.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | slope | depth available to roots | acidity (pH) | texture | erosion | stoniness | soil organic matter | water capacity | flood | sodicity | clay | sand | contamination | cation exchange capacity | productivity | dryness | natural toxicity |
| slope | 1.00 | 1.00 | 2.00 | 2.00 | 3.00 | 3.00 | 3.00 | 4.00 | 4.00 | 5.00 | 6.00 | 6.00 | 6.00 | 6.00 | 7.00 | 8.00 | 9.00 |
| depth available to roots | 1.00 | 1.00 | 2.00 | 2.00 | 3.00 | 3.00 | 3.00 | 4.00 | 4.00 | 5.00 | 6.00 | 6.00 | 6.00 | 6.00 | 7.00 | 8.00 | 9.00 |
| acidity (pH) | 0.50 | 0.50 | 1.00 | 1.00 | 1.50 | 1.50 | 1.50 | 2.00 | 2.00 | 2.50 | 3.00 | 3.00 | 3.00 | 3.00 | 3.50 | 4.00 | 4.50 |
| texture | 0.50 | 0.50 | 1.00 | 1.00 | 1.50 | 1.50 | 1.50 | 2.00 | 2.00 | 2.50 | 3.00 | 3.00 | 3.00 | 3.00 | 3.50 | 4.00 | 4.50 |
| erosion | 0.33 | 0.33 | 0.67 | 0.67 | 1.00 | 1.00 | 1.00 | 1.33 | 1.33 | 1.67 | 2.00 | 2.00 | 2.00 | 2.00 | 2.33 | 2.67 | 3.00 |
| stoniness | 0.33 | 0.33 | 0.67 | 0.67 | 1.00 | 1.00 | 1.00 | 1.33 | 1.33 | 1.67 | 2.00 | 2.00 | 2.00 | 2.00 | 2.33 | 2.67 | 3.00 |
| soil organic matter | 0.33 | 0.33 | 0.67 | 0.67 | 1.00 | 1.00 | 1.00 | 1.33 | 1.33 | 1.67 | 2.00 | 2.00 | 2.00 | 2.00 | 2.33 | 2.67 | 3.00 |
| water capacity | 0.33 | 0.33 | 0.67 | 0.67 | 1.00 | 1.00 | 1.00 | 1.33 | 1.33 | 1.67 | 2.00 | 2.00 | 2.00 | 2.00 | 2.33 | 2.67 | 2.25 |
| flood | 0.25 | 0.25 | 0.50 | 0.50 | 0.75 | 0.75 | 0.75 | 1.00 | 1.00 | 1.25 | 1.50 | 1.50 | 1.50 | 1.50 | 1.75 | 2.00 | 2.25 |
| sodicity | 0.25 | 0.25 | 0.50 | 0.50 | 0.75 | 0.75 | 0.75 | 1.00 | 1.00 | 1.25 | 1.50 | 1.50 | 1.50 | 1.50 | 1.75 | 2.00 | 1.80 |
| clay | 0.20 | 0.20 | 0.40 | 0.40 | 0.60 | 0.60 | 0.60 | 0.80 | 0.80 | 1.00 | 1.20 | 1.20 | 1.20 | 1.20 | 1.40 | 1.60 | 1.50 |
| sand | 0.17 | 0.17 | 0.33 | 0.33 | 0.50 | 0.50 | 0.50 | 0.67 | 0.67 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.17 | 1.33 | 1.50 |
| contamination | 0.17 | 0.17 | 0.33 | 0.33 | 0.50 | 0.50 | 0.50 | 0.67 | 0.67 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.17 | 1.33 | 1.50 |
| cation exchange capacity | 0.17 | 0.17 | 0.33 | 0.33 | 0.50 | 0.50 | 0.50 | 0.67 | 0.67 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.17 | 1.33 | 1.50 |
| productivity | 0.17 | 0.17 | 0.33 | 0.33 | 0.50 | 0.50 | 0.50 | 0.67 | 0.67 | 0.83 | 1.00 | 1.00 | 1.00 | 1.00 | 1.17 | 1.33 | 1.29 |
| dryness | 0.14 | 0.14 | 0.29 | 0.29 | 0.43 | 0.43 | 0.43 | 0.57 | 0.57 | 0.71 | 0.86 | 0.86 | 0.86 | 0.86 | 1.00 | 1.14 | 1.13 |
| natural toxicity | 0.13 | 0.13 | 0.25 | 0.25 | 0.38 | 0.38 | 0.38 | 0.50 | 0.50 | 0.63 | 0.75 | 0.75 | 0.75 | 0.75 | 0.88 | 1.00 | 1.00 |

Table 20: The normalized pairwise comparison matrix and the calculated weights.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | slope | depth available to roots | acidity (pH) | texture | erosion | stoniness | soil organic matter | water capacity | flood | sodicity | clay | sand | contamination | cation exchange capacity | productivity | dryness | natural toxicity | **Weights** |
| slope | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.22 | 0.17 | 0.21 | 0.20 | 0.17 | 0.17 | 0.17 | 0.20 | 0.19 | 0.19 | 0.18 |
| depth available to roots | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.17 | 0.22 | 0.17 | 0.21 | 0.20 | 0.17 | 0.17 | 0.17 | 0.20 | 0.19 | 0.19 | 0.18 |
| acidity (pH) | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.11 | 0.08 | 0.10 | 0.10 | 0.08 | 0.08 | 0.08 | 0.10 | 0.10 | 0.09 | 0.09 |
| texture | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.11 | 0.08 | 0.10 | 0.10 | 0.08 | 0.08 | 0.08 | 0.10 | 0.10 | 0.09 | 0.09 |
| erosion | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.07 | 0.07 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.06 | 0.06 |
| stoniness | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.07 | 0.07 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.06 | 0.06 |
| soil organic matter | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.07 | 0.07 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.06 | 0.06 |
| water capacity | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.07 | 0.07 | 0.06 | 0.06 | 0.06 | 0.07 | 0.06 | 0.05 | 0.06 |
| flood | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.04 |
| sodicity | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.04 | 0.04 |
| clay | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.04 | 0.04 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 | 0.03 | 0.04 |
| sand | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| contamination | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| cation exchange capacity | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| productivity | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| dryness | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.02 | 0.03 |
| natural toxicity | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |

Table 21: Terrain and Soil indicator classes, scores and weights.

|  |  |  |  |
| --- | --- | --- | --- |
| **Indicator** | **Classes** | **Score** | **Weight** |
| slope | [15% - 40%]  [40% - 65%]  [65% - 90%] | 10  5  1 | 0.18 |
| depth available to roots | [100cm – 66.7cm]  [66.7cm – 33.3cm]  [33.3cm – 0cm] | 10  5  1 | 0.18 |
| texture | [30% - 53.3%]  [53.3% - 76.7%]  [76.7% - 100%] | 10  5  1 | 0.09 |
| stoniness | [10% - 15%]  [15% - 20%] | 10  5 | 0.06 |
| water capacity | [100mm – 50mm]  [50mm – 0mm] | 10  5 | 0.06 |
| clay | [50% - 58.7%]  [58.7% - 67.3%]  [67.3% - 76%] | 10  5  1 | 0.04 |
| sand | [60% - 70%]  [70% - 80%]  [80% - 90%] | 10  5  1 | 0.03 |

Table 22: Sustainability indicator classes, scores and weights.

|  |  |  |  |
| --- | --- | --- | --- |
| **Indicator** | **Classes** | **Score** | **Weight** |
| acidity (pH) | [pH>8, pH<6]  [pH>8.5, pH<5.25]  [pH>9, pH<4.5] | 10  5  1 | 0.09 |
| erosion | [200 – 241.7]  [241.7 – 283.4]  [283.4- 325] | 10  5  1 | 0.06 |
| flood | [50% - 66.7%]  [66.7% - 83.3%]  [83.3% - 100%] | 10  5  1 | 0.04 |
| sodicity | [6% - 36.7%]  [36.7% - 67.4%]  [67.4% - 98%] | 10  5  1 | 0.04 |
| contamination | [1cg/kg – 3cg/kg]  [3cg/kg – 10cg/kg]  [10cg/kg – 23.5cg/kg] | 10  5  1 | 0.03 |
| dryness | [0.5 – 0.34]  [0.34 – 0.18]  [0.18 – 0] | 10  5  1 | 0.03 |
| natural toxicity | [150g/Kg - 328g/Kg]  [328g/Kg - 506g/Kg]  [506g/Kg – 684g/Kg] | 10  5  1 | 0.02 |

Table 23: Productivity indicator classes, scores and weights.

|  |  |  |  |
| --- | --- | --- | --- |
| **Indicator** | **Classes** | **Score** | **Weight** |
| soil organic matter | [OM < 1%, OM ≥ 20%]  [OM < 0.75%, OM ≥ 30%] | 10  5 | 0.06 |
| cation exchange capacity | [8 – 5.3]  [5.3 – 2.7]  [2.7 – 0] | 10  5  1 | 0.03 |
| productivity | Grasslands:  [6-4]  [4-2]  [2-0]  Forests:  [3-2]  [2-1]  [1-0] | 10  5  1  10  5  1 | 0.03 |

## Production of intermediate layers

Some of the indicators used in the analysis, namely *soil texture* and *dryness*, are calculated based on other variables. These indicators will be referred to as “synthetic” indicators and require the production of intermediate layers in the GIS modelling procedure.

More specifically, the *soil texture* layer consists of the *stoniness, clay* and *sand* variables (all with independent thresholds), along with an intermediate layer calculated as (Elbersen et al.) :

Where *silt* and *clay* are percentages found in soil. The threshold applied to this intermediate layer is ≥ 30%, for marginal lands.

Moreover, the *dryness* layer is created by computing the Aridity Index (AI), which consists of *precipitation* and *evapotranspiration* and is calculated as:

Where *P* is the total annual precipitation and *PET* is the total annual potential evapotranspiration.

Apart from the indicator intermediate layers, another intermediate layer that needs to be calculated is the *weighted overlay (WO)* layer that will be:

Where *S* are the scores assigned for each indicator class and *W* are the weights assigned for each indicator (according to Table 21, Table 22 and Table 23).

# Development and realization of a GIS using the datasets collected in T2.2

## Datasets overview

The 201 datasets collected in T2.2, were filtered based on the indicators found in the literature review. Datasets that were closest to the indicators conceptually were chosen, resulting initially in a sub-collection of 70 datasets (Datasets v1). These datasets were then filtered again, and by concurrently examining the data and thresholds used in literature, a best match was found, resulting in the final datasets (Datasets v2) and threshold simultaneously. The Datasets v1 can be thus used as a reference for potential changes in the methodology and thresholds used.

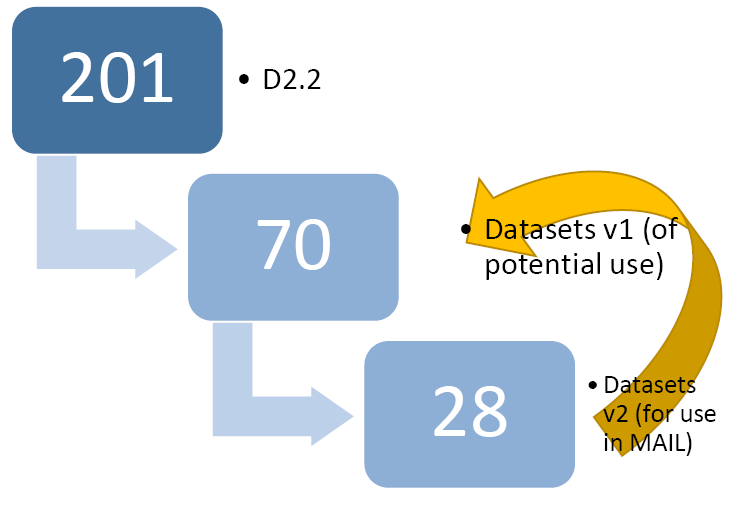


Figure 6: The elimination steps followed for selecting the final datasets used in *MAIL*.

Table 24: Final datasets used in *MAIL*.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Soil** | **Climate** | **Terrain** | **Sustainability** | **Productivity** | **LULC** | **Raster/Vector** | **CRS** | **Spatial Resolution / MMU** | **Coverage** | **Product Date** | **Values Range** |
|
| Total available water content from PTF |  |  |  |  |  | R | ETRS89 | 1km | european | 2013 |  |
| Clay content (topsoil & subsoil) |  |  |  |  |  | R | ETRS89 | 1km | european | 2013 | 0-76 |
| Depth available to roots |  |  |  |  |  | R | ETRS89 | 1km | european | 2013 | 0-150 |
| Sand content (topsoil & subsoil) |  |  |  |  |  | R | ETRS89 | 1km | european | 2013 | 0-90 |
| Silt content (topsoil & subsoil) |  |  |  |  |  | R | ETRS89 | 1km | european | 2013 | 0-71 |
| Soil pH in Europe |  |  |  |  |  | R | ETRS89 | 5km | EU25 (Romania & Bulgaria are not included,)+Norway, Switzerland, Croatia, Albania | 2009 | 0.6-8.8 |
| Volume of stones |  |  |  |  |  | V | WGS84 |  | EU27 | 2008 | 0%-20% |
| WRB-FULL. Full soil code of the STU from the World Reference Base (WRB) for Soil Resources |  |  |  |  |  | V | WGS84 |  | EU27 | 2008 | thematic |
|  |  | EU-DEM |  |  |  | R | ETRS89 | 25m | european | 2011 |  |
|  |  |  | Soil erosion by water (RUSLE2015) |  |  | R | ETRS89 | 100m | EU28 | 2015 | 0-325 |
|  | TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces |  | TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces |  |  | R | WGS84 | ~4.5km | global | 1958-2018 | PET: 80-239, PC: 0-1559 |
| WISE derived soil property estimates |  |  | WISE derived soil property estimates |  |  | R | WGS84 | ~10km | global | 2015 | sodicity: ESP [-9-98] , nitrogen: TOTN [-9-23.48] , gypsum: GYPS [-9-684] , CEC: CECs [-9-128], soil types |
|  |  |  |  | Organic carbon content (topsoil & subsoil) |  | R | ETRS89 | 1km | european | 2013 | 0-39.5 |
|  |  |  |  | Soil biomass productivity of forest areas |  | R | ETRS89 | 1km | EU27 | 2016 | 0.17-10 |
|  |  |  |  | Soil biomass productivity of grasslands and pastures |  | R | ETRS89 | 1km | EU27 | 2016 | 0-10 |
|  |  |  |  |  | Coastline EU | V | ETRS89 |  | european |  | thematic |
|  |  |  |  |  | CORINE LC | V | ETRS89 | 25ha | EEA39 | 2018 | thematic |
|  |  |  |  |  | CORINE LC change | V | ETRS89 | 5ha | EEA39 | 2018 | thematic |
|  |  |  |  |  | Global Forest Change | R | WGS84 | 30m | global | 2000-2018 | thematic |
|  |  |  |  |  | Imperviousness Classified Change (IMCC) | R | ETRS89 | 20m | european | 2012-2015 | thematic |
|  |  |  |  |  | Imperviousness Density (IMD) HRL | R | ETRS89 | 20m | european | 2015 | thematic |
|  |  |  | JRC Global Surface Water |  | JRC Global Surface Water | R | WGS84 | 30m | global | 1984-2019 | 0-100% |
|  |  |  |  |  | Nationally designated areas (CDDA) | V | WGS84 |  | european | 2019 | thematic |
|  |  |  |  |  | Natura2000 | V | ETRS89 |  | EEA33 | 2018 | thematic |
|  |  |  |  |  | S2GLC | R | ETRS89 | 10m-60m | european | 2017 | thematic |
|  |  |  |  |  | TanDEM-X Global Forest map | R | WGS84 | 50m | global | 2011-2015 | thematic |
|  |  |  |  |  | Tree Cover Density (TCD) | R | ETRS89 | 20m | european | 2015 | thematic |
|  |  |  |  |  | Tree Cover Density Change (TCDC) | R | ETRS89 | 20m | european | 2012-2015 | 1-100% |

In addition to datasets from T2.2, three new datasets were added to cover the needs for two indicators, namely the “Global Surface Water”[[14]](#footnote-14) dataset by the Joint Research Center, the “Global Forest Change”[[15]](#footnote-15) dataset by the University of Maryland and the “TerraClimate”[[16]](#footnote-16) dataset, by the University of Idaho (Table 25). The “Global Surface Water” dataset is used calculating the flooding/ponding indicator and the “Global Forest Change” dataset is used as a land cover dataset, for forest areas detection and masking. Moreover, based on the datasets used in the studied literature, 7 more datasets are suggested to MAIL, for potential use in the future (Table 26) and are included in Datasets v1. It can be noticed in Table 24, that the coverage of each dataset differs. This will cause issues in the workflow implementation and final MLs detection, since the indicators with missing data will not contribute to the MLs detection in the area where missing. In these cases, the workflow will continue by omitting the corresponding step of this indicator, and a flag will be assigned to the areas/pixels where calculation was done with missing values.

Table 25: Additional datasets for use in *MAIL*.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Dataset Name** | **Raster/Vector** | **CRS** | **Spatial Resolution/MMU** | **Coverage** | **Product Date** | **Use** |
| Global Surface Water | R | WGS84 | 30m | global | 1984-2019 | flooding indicator |
| Global Forest Change | R | WGS84 | 30m | global | 2000-2018 | forest mask |
| TerraClimate | R |  | 2.5 arc minutes | global | 1958-2015 | aridity index (precipitation accumulation, evapotranspiration) |

Table 26: Datasets for potential use in MAIL, as an addition to datasets from T2.2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Dataset Name** | **Raster/Vector** | **CRS** | **Spatial Resolution/MMU** | **Coverage** | **Product Date** | **Potential Use** |
| SoilGrids | R | WGS84 | 250m | global | 2016 | acidity, cation exchange capacity, texture |
| MODIS LST | R | Sinusoidal | 1km | global | 2007-2019 | Land surface temperature |
| MODIS NPP | R | Sinusoidal | 1km | global | 2000-2014 | productivity |
| TRMM 3B43 | R |  | 0.25 arc degrees | global | 1998-2019 | monthly and annual precipitation |
| GHSL population | R | WGS84 | 250m | global | 2015 | population / socioeconomic indicators |
| K1 mountain | R | WGS84 | 1km | global | 2000 (1996) |  |
| K3 mountain | R | WGS84 | 250m | global | 2017 (2010) |  |

## Dataset transformation to common file formats, projection, etc

Most of the final datasets in Datasets v2 (Table 24) were in the European Terrestrial Reference System 1989 (ETRS89), whereas others where in the World Geodetic System 1984 (WGS84). Since it was decided to use the ETRS89 as a coordinate system for all layers in MAIL, the datasets in WGS84 were transformed using the ArcGIS “Project Raster” tool, into ETRS89.

Resolution of all datasets varies from 10m (S2GLC) to 10km (WISE derived soil property estimates dataset). Layers’ original resolution will be kept as is, during the threshold phase of the workflow and only before the weighted overlay phase, a unique cell size equal to the finest dataset resolution (25m – EU DEM) will be applied to the data used for ranking. Data will be resampled to the marginal land MMU (1ha) in the end of all processes.

## GIS implementation

All datasets (vector and raster) that were collected based on the methodology development (Datasets v2), were imported into a geodatabase and then into a GIS project. There, the datasets were transformed to the ETRS89 coordinate system and the layers created, were grouped into categories (soil, climate, terrain, sustainability, productivity, LULC, socioeconomic) (Figure 7). This initial GIS project will work as a base for the workflows implemented in the next stage of the project.

The GIS software which was used in this phase and will be also used in the next steps, is the ArcGIS[[17]](#footnote-17) software by Esri, with a license provided by IABG.

Due to the large scale of the study and to the large datasets, tile structure is used for the raster data in the geodatabase. Raster datasets were imported in the form of Raster Catalogs[[18]](#footnote-18) and Raster Mosaics[[19]](#footnote-19). A raster catalog is a simple container for managing raster datasets, whereas a mosaic dataset is more advanced as overview images can be built and a processing for each raster dataset or on the entire mosaic dataset can be defined. Both are stored within a geodatabase.

Additionally, tiled processing components and parallel processing will be used during the workflow run. To improve the performance and scalability of feature overlay tools, operational logic called *adaptive subdivision processing*[[20]](#footnote-20) will be used. This is a built-in feature of ArcGIS that is enabled when data cannot be processed within the available amount of physical memory, and thus, all data are subdivided in tiles, processed separately and then reconstructed again.

Finally, it is important to stress out that, the workflows implementing the methodology developed, should have an independency in structure, so that potential changes in one step will not affect the next steps in processing. This is most important to be done in the ranking, scoring and weighted overlay phase, which is most prone to changing.

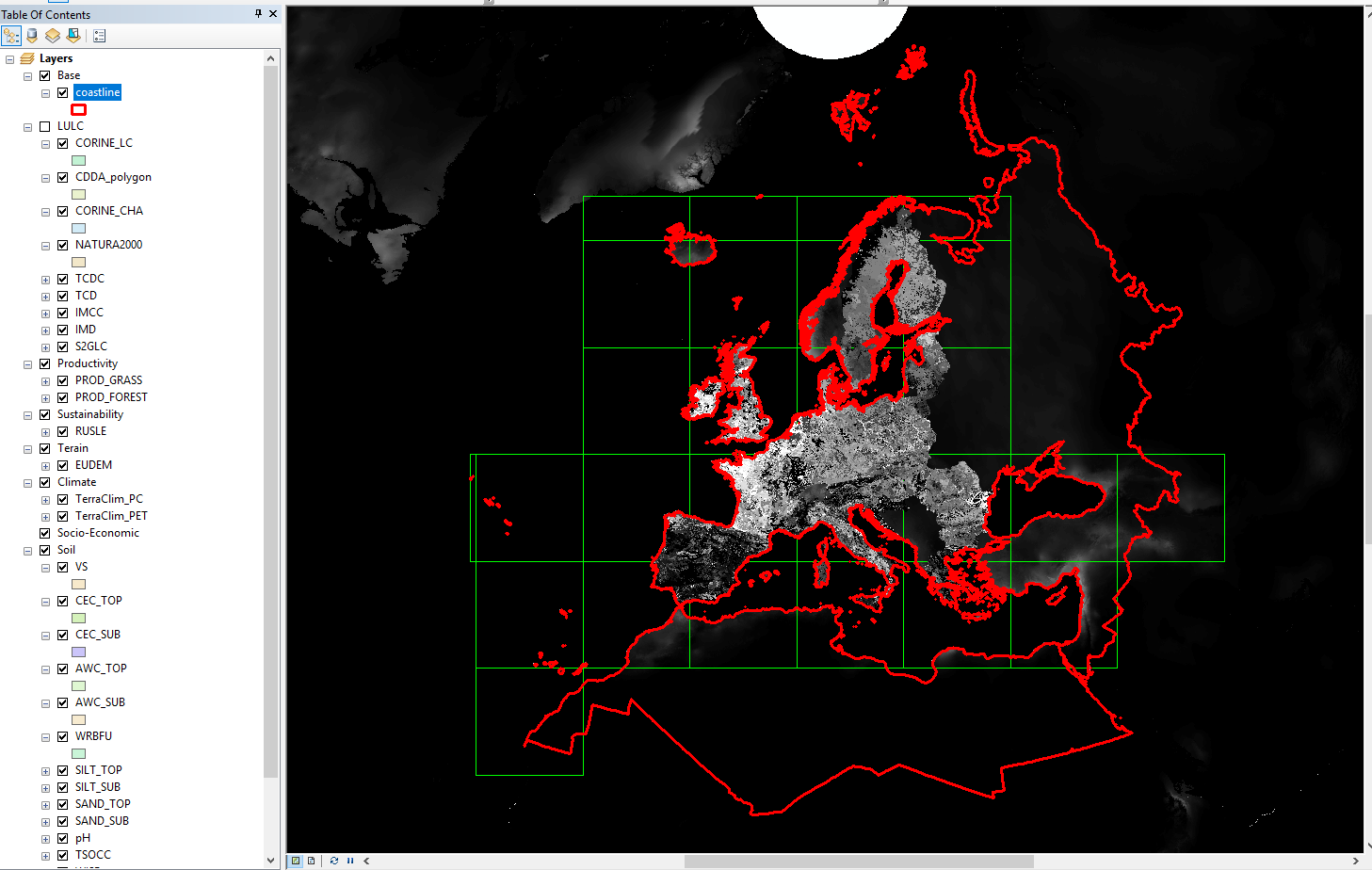


Figure 7: Datasets and layers as seen in the ArcGIS project.

## Layers

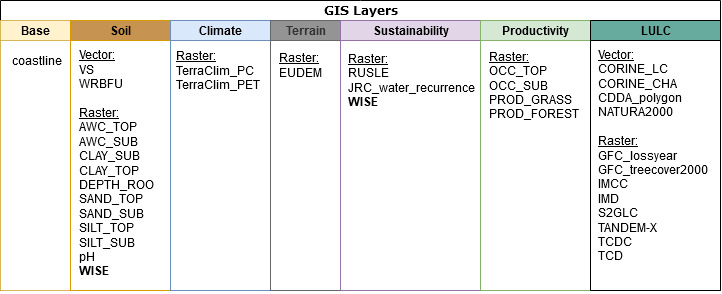


Figure 8: The final names of the Datasets v2 GIS layers.

# Evaluation of existing and development of indicator sets

## Selection of test sites

## Testing different sets of existing indicators and thresholds

## Development of new indicators sets

## Refinement of indicators and criteria

# Design and development of computation workflows.

## Workflow design for refined indicator sets

## Algorithm development for the production of new layers incorporating initial datasets

# Implementation of workflows to the GIS

## Workflow programming

## Automatization of procedures

## Workflow testing

## Produced tools description

# Development of guidelines for the usage indicators set and GIS workflows

# Experimental results for MLs identification

# 1st methodology refinement (indicator fine tuning based on the experimental results

# Refinement of results based on Τ2.4 results T2.8 recommendations

# References

1. Bertaglia, M., Joost, S., & Roosen, J. (2007). Identifying European marginal areas in the context of local sheep and goat breeds conservation: A geographic information system approach. Agricultural Systems, 94(3), 657–670. <https://doi.org/10.1016/j.agsy.2007.02.006>
2. Büttner, G., & Kosztra, B. (2017). CLC2018 technical guidelines. European Environment Agency, Wien. Retrieved November 25, 2019, from <https://land.copernicus.eu/user-corner/technical-library/clc2018technicalguidelines_final.pdf>
3. Chen, J., Yang, S., Li, H. W., Zhang, B., & Lv, J. (2013). Research on geographical environment unit division based on the method of natural breaks (Jenks). Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, 47-50.
4. Copernicus, High Resolution Layers, HRL. (2012, 2015). Retrieved August 08, 2019, from <https://land.copernicus.eu/pan-european/high-resolution-layers>
5. Dzieszko, P. (2014). Land-cover modelling using corine land cover data and multi-layer perceptron. Quaestiones Geographicae, 33(1), 5-22.
6. Grassi, G., Cescatti, A., Matthews, R., Duveiller, G., Camia, A., Federici, S., Vizzarri, M. (2019). On the realistic contribution of European forests to reach climate objectives. Carbon Balance and Management, 14(1), 8. <https://doi.org/10.1186/s13021-019-0123-y>
7. Hamad, R., Balzter, H., & Kolo, K. (2018). Predicting land use/land cover changes using a CA-Markov model under two different scenarios. Sustainability, 10(10), 3421.
8. European Environment Agency, Digital Elevation Model over Europe (EU- DEM). (2017, December 7). Retrieved July 29, 2019, from <http://www.eea.europa.eu/data-and-maps/data/eu-dem#tab-european-data>
9. European Environment Agency, Natura 2000 data - the European network of Protected sites. (2019, April 12). Retrieved July 17, 2019, from <https://www.eea.europa.eu/data-and-maps/data/natura-10#tab-additional-information>
10. European Parliament. Regulation (EU) 2018/841 of the European Parliament and of the Council of 30 May 2018 on the inclusion of greenhouse gas emissions and removals from land use, land use change and forestry in the 2030 climate and energy framework., 2018 Official Journal of the European Union § (2018)
11. Romppainen, S. (2019). How does the LULUCF Regulation affect EU Member States’ forest management? Govrtran Policy Brief, 03(January 2019).
12. Tutić, D., Štanfel, M., & Horvat, M. (2018). Multi-Criteria Land Evaluation of Suitability for the Sport of Foot Orienteering: A Case Study of Croatia and Slovenia. ISPRS International Journal of Geo-Information, 7(6), 227

Van Echelpoel, W., Boets, P., Landuyt, D., Gobeyn, S., Everaert, G., Bennetsen, E., ... & Goethals, P. L. (2015). Species distribution models for sustainable ecosystem management. In Developments in Environmental Modelling (Vol. 27, pp. 115-134). Elsevier.

Bai, Z., Dent, D., Olsson, L., & Schaepman, M. (2008). Global Assessment of Land Degradation and Improvement 1. Identification by remote sensing. *ISRIC – World Soil Information, Wageningen*, (November), 69. https://doi.org/report 2008/01

Bandaru, V., Izaurralde, R. C., Manowitz, D., Link, R., Zhang, X., & Post, W. M. (2013). Soil Carbon Change and Net Energy Associated with Biofuel Production on Marginal Lands: A Regional Modeling Perspective. *Journal of Environment Quality*, *42*(6), 1802. https://doi.org/10.2134/jeq2013.05.0171

Bertaglia, M., Joost, S., & Roosen, J. (2007). Identifying European marginal areas in the context of local sheep and goat breeds conservation: A geographic information system approach. *Agricultural Systems*, *94*(3), 657–670. https://doi.org/10.1016/j.agsy.2007.02.006

Cai, X., Zhang, X., & Wang, D. (2011). Land Availability for Biofuel Production. *Environmental Science & Technology*, *45*(1), 334–339. https://doi.org/10.1021/es103338e

Ciria, C., Sanz, M., Carrasco, J., & Ciria, P. (2019). Identification of Arable Marginal Lands under Rainfed Conditions for Bioenergy Purposes in Spain. *Sustainability*, *11*(7), 1833. https://doi.org/10.3390/su11071833

Clark, D. A., Brown, S., Kicklighter, D. W., Chambers, J. Q., Thomlinson, J. R., & Ni, J. (2001). Measuring net primary production in forest: concepts and field methods. *Ecological Society of America*, *11*(2), 356–370.

Elbersen, B., Eupen, van E., Mantel, S., Verzandvoort, S., Boogaard, H., Mucher, S., … Eleftheriadis, I. (n.d.). *Deliverable 2.6 Methodological approaches to identify and map marginal land suitable for industrial crops in Europe*.

Eliasson, Å., Jones, R. J. A., Nachtergaele, F., Rossiter, D. G., Terres, J. M., van Orshoven, J., … Le Bas, C. (2010). Common criteria for the redefinition of Intermediate Less Favoured Areas in the European Union. *Environmental Science and Policy*, *13*(8), 766–777. https://doi.org/10.1016/j.envsci.2010.08.003

Gelfand, I., Sahajpal, R., Zhang, X., Izaurralde, R. C., Gross, K. L., & Robertson, G. P. (2013). Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*, *493*(7433), 514–517. https://doi.org/10.1038/nature11811

Gerwin, W., Repmann, F., Galatsidas, S., Vlachaki, D., Gounaris, N., Baumgarten, W., … Freese, D. (2018). Assessment and quantification of marginal lands for biomass production in Europe using soil-quality indicators. *Soil*, *4*(4), 267–290. https://doi.org/10.5194/soil-4-267-2018

Gopalakrishnan, G., Cristina Negri, M., & Snyder, S. W. (2011). A Novel Framework to Classify Marginal Land for Sustainable Biomass Feedstock Production. *Journal of Environment Quality*, *40*(5), 1593. https://doi.org/10.2134/jeq2010.0539

Ivanina, V., Roik, M., & Hanzhenko, O. (2016). *Report on MagL concepts, debate and indicators (D2.3)*. *D2.3 of SEEMLA project*. Ukraine. Retrieved from http://www.seemla.eu/en/home/

JRC. (2018). *Global Surface Water - Data Users Guide (v2) Background*. Retrieved from https://global-surface-water.appspot.com/download

Kang, S., Post, W. M., Wang, D., Nichols, J. A., Bandaru, V., & West, T. O. (2013). Hierarchical marginal land assessment for land use planning. *Land Use Policy*, *30*(1), 106–113. https://doi.org/10.1016/j.landusepol.2012.03.002

Li, G., Messina, J. P., Peter, B. G., & Snapp, S. S. (2017). Mapping Land Suitability for Agriculture in Malawi. *Land Degradation and Development*, *28*(7), 2001–2016. https://doi.org/10.1002/ldr.2723

Milbrandt, A., & Overend, R. P. (2009). Assessment of Biomass Resources from Marginal Lands in APEC Economies. In *APEC Energy Working Group* (p. 52).

Mueller, L., Schindler, U., Behrendt, A., Eulenstein, F., & Dannowski, R. (2007). *The Muencheberg Soil Quality Rating (SQR): FIELD MANUAL FOR DETECTING AND ASSESSING PROPERTIES AND LIMITATIONS OF SOILS FOR CROPPING AND GRAZING* (Vol. 2014). Muencheberg, Germany. Retrieved from http://www.zalf.de/en/forschung/institute/lwh/mitarbeiter/lmueller/Documents/field\_mueller.pdf

Niu, X., & Duiker, S. W. (2006). Carbon sequestration potential by afforestation of marginal agricultural land in the Midwestern U.S. *Forest Ecology and Management*, *223*(1–3), 415–427. https://doi.org/10.1016/j.foreco.2005.12.044

Peter, B. G., Messina, J. P., & Snapp, S. S. (2018). Multiscalar approach to mapping marginal agricultural land: Smallholder agriculture in Malawi. *Annals of the American Association of Geographers*, *108*(4), 989–1005. https://doi.org/10.1080/24694452.2017.1403877

Roehrig, J., & Menz, G. (2005). The Determination of Natural Agricultural Potential in Western Africa Using the Fuzzy Logic Based Marginality Index. *EARSeL EProceedings*, 9–17.

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*(February), 526–534. https://doi.org/10.1016/j.landusepol.2018.02.033

Schweers, W., Campbell, E., Hennenberg, K., Fritsche, U., Mang, H. P., Lucas, M., … Zhang, N. (2011). Identification of potential areas for biomass production in China: Discussion of a recent approach and future challenges. *Biomass and Bioenergy*, *35*(5), 2268–2279. https://doi.org/10.1016/j.biombioe.2011.02.034

Zolekar, R. B. (2018). Integrative approach of RS and GIS in characterization of land suitability for agriculture: a case study of Darna catchment. *Arabian Journal of Geosciences*, *11*(24). https://doi.org/10.1007/s12517-018-4148-4

Zolekar, R. B., & Bhagat, V. S. (2015). Multi-criteria land suitability analysis for agriculture in hilly zone: Remote sensing and GIS approach. *Computers and Electronics in Agriculture*, *118*, 300–321. https://doi.org/10.1016/j.compag.2015.09.016

Bai, Z., Dent, D., Olsson, L., & Schaepman, M. (2008). Global Assessment of Land Degradation and Improvement 1. Identification by remote sensing. *ISRIC – World Soil Information, Wageningen*, (November), 69. https://doi.org/report 2008/01

Bandaru, V., Izaurralde, R. C., Manowitz, D., Link, R., Zhang, X., & Post, W. M. (2013). Soil Carbon Change and Net Energy Associated with Biofuel Production on Marginal Lands: A Regional Modeling Perspective. *Journal of Environment Quality*, *42*(6), 1802. https://doi.org/10.2134/jeq2013.05.0171

Bertaglia, M., Joost, S., & Roosen, J. (2007). Identifying European marginal areas in the context of local sheep and goat breeds conservation: A geographic information system approach. *Agricultural Systems*, *94*(3), 657–670. https://doi.org/10.1016/j.agsy.2007.02.006

Cai, X., Zhang, X., & Wang, D. (2011). Land Availability for Biofuel Production. *Environmental Science & Technology*, *45*(1), 334–339. https://doi.org/10.1021/es103338e

Ciria, C., Sanz, M., Carrasco, J., & Ciria, P. (2019). Identification of Arable Marginal Lands under Rainfed Conditions for Bioenergy Purposes in Spain. *Sustainability*, *11*(7), 1833. https://doi.org/10.3390/su11071833

Clark, D. A., Brown, S., Kicklighter, D. W., Chambers, J. Q., Thomlinson, J. R., & Ni, J. (2001). Measuring net primary production in forest: concepts and field methods. *Ecological Society of America*, *11*(2), 356–370.

Elbersen, B., Eupen, van E., Mantel, S., Verzandvoort, S., Boogaard, H., Mucher, S., … Eleftheriadis, I. (n.d.). *Deliverable 2.6 Methodological approaches to identify and map marginal land suitable for industrial crops in Europe*.

Eliasson, Å., Jones, R. J. A., Nachtergaele, F., Rossiter, D. G., Terres, J. M., van Orshoven, J., … Le Bas, C. (2010). Common criteria for the redefinition of Intermediate Less Favoured Areas in the European Union. *Environmental Science and Policy*, *13*(8), 766–777. https://doi.org/10.1016/j.envsci.2010.08.003

Gelfand, I., Sahajpal, R., Zhang, X., Izaurralde, R. C., Gross, K. L., & Robertson, G. P. (2013). Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*, *493*(7433), 514–517. https://doi.org/10.1038/nature11811

Gerwin, W., Repmann, F., Galatsidas, S., Vlachaki, D., Gounaris, N., Baumgarten, W., … Freese, D. (2018). Assessment and quantification of marginal lands for biomass production in Europe using soil-quality indicators. *Soil*, *4*(4), 267–290. https://doi.org/10.5194/soil-4-267-2018

Gopalakrishnan, G., Cristina Negri, M., & Snyder, S. W. (2011). A Novel Framework to Classify Marginal Land for Sustainable Biomass Feedstock Production. *Journal of Environment Quality*, *40*(5), 1593. https://doi.org/10.2134/jeq2010.0539

Ivanina, V., Roik, M., & Hanzhenko, O. (2016). *Report on MagL concepts, debate and indicators (D2.3)*. *D2.3 of SEEMLA project*. Ukraine. Retrieved from http://www.seemla.eu/en/home/

JRC. (2018). *Global Surface Water - Data Users Guide (v2) Background*. Retrieved from https://global-surface-water.appspot.com/download

Kang, S., Post, W. M., Wang, D., Nichols, J. A., Bandaru, V., & West, T. O. (2013). Hierarchical marginal land assessment for land use planning. *Land Use Policy*, *30*(1), 106–113. https://doi.org/10.1016/j.landusepol.2012.03.002

Li, G., Messina, J. P., Peter, B. G., & Snapp, S. S. (2017). Mapping Land Suitability for Agriculture in Malawi. *Land Degradation and Development*, *28*(7), 2001–2016. https://doi.org/10.1002/ldr.2723

Milbrandt, A., & Overend, R. P. (2009). Assessment of Biomass Resources from Marginal Lands in APEC Economies. In *APEC Energy Working Group* (p. 52).

Mueller, L., Schindler, U., Behrendt, A., Eulenstein, F., & Dannowski, R. (2007). *The Muencheberg Soil Quality Rating (SQR): FIELD MANUAL FOR DETECTING AND ASSESSING PROPERTIES AND LIMITATIONS OF SOILS FOR CROPPING AND GRAZING* (Vol. 2014). Muencheberg, Germany. Retrieved from http://www.zalf.de/en/forschung/institute/lwh/mitarbeiter/lmueller/Documents/field\_mueller.pdf

Niu, X., & Duiker, S. W. (2006). Carbon sequestration potential by afforestation of marginal agricultural land in the Midwestern U.S. *Forest Ecology and Management*, *223*(1–3), 415–427. https://doi.org/10.1016/j.foreco.2005.12.044

Peter, B. G., Messina, J. P., & Snapp, S. S. (2018). Multiscalar approach to mapping marginal agricultural land: Smallholder agriculture in Malawi. *Annals of the American Association of Geographers*, *108*(4), 989–1005. https://doi.org/10.1080/24694452.2017.1403877

Roehrig, J., & Menz, G. (2005). The Determination of Natural Agricultural Potential in Western Africa Using the Fuzzy Logic Based Marginality Index. *EARSeL EProceedings*, 9–17.

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*(February), 526–534. https://doi.org/10.1016/j.landusepol.2018.02.033

Schweers, W., Campbell, E., Hennenberg, K., Fritsche, U., Mang, H. P., Lucas, M., … Zhang, N. (2011). Identification of potential areas for biomass production in China: Discussion of a recent approach and future challenges. *Biomass and Bioenergy*, *35*(5), 2268–2279. https://doi.org/10.1016/j.biombioe.2011.02.034

Zolekar, R. B. (2018). Integrative approach of RS and GIS in characterization of land suitability for agriculture: a case study of Darna catchment. *Arabian Journal of Geosciences*, *11*(24). https://doi.org/10.1007/s12517-018-4148-4

Zolekar, R. B., & Bhagat, V. S. (2015). Multi-criteria land suitability analysis for agriculture in hilly zone: Remote sensing and GIS approach. *Computers and Electronics in Agriculture*, *118*, 300–321. https://doi.org/10.1016/j.compag.2015.09.016

# Annex I: Marginality index calculation and CLC reclassification

Land use class chosen[[21]](#footnote-21) for less productive and non-productive land is the class 3.2 (‘‘Low vegetation”). Class 1.1.1 (“Artificial surfaces and constructions”) and class 2.1 (“Cultivated and managed areas”) are used as an indicator of a less marginal land use. Thus, the variable computed is the ratio of the percent of the neighborhood area (for further explanation see below) classified as land use class 3.2 to the percent of area in land use class 1.1.1 and class 2.1, that is:

MARINDEX = (%class 3.2) / (%class1.1.1 + %class2.1)

The standard CLC nomenclature includes 44 land cover classes[[22]](#footnote-22). These are grouped in a three-level hierarchy. The five main (level-one) categories are: 1) artificial surfaces, 2) agricultural areas, 3) forests and semi-natural areas, 4) wetlands, 5) water bodies. All national teams had to adopt this standard nomenclature according to their landscape conditions. Although the 44 categories have not changed since the implementation of the first CLC inventory (1986-1998), the definition of most of the nomenclature elements was significantly improved.

| **CLC Level 1** | **CLC Level 2** | **CLC Level 3** |
| --- | --- | --- |
| 1 Artificial surfaces | 1.1 Urban fabric | 1.1.1 Continuous urban fabric |
| 1.1.2 Discontinuous urban fabric |
| 1.2 Industrial, commercial and transport units | 1.2.1 Industrial and commercial units |
| 1.2.2 Road and rail networks and associated land |
| 1.2.3 Port areas |
| 1.2.4 Airports |
| 1.3 Mine, dump and construction sites | 1.3.1 Mineral extraction sites |
| 1.3.2 Dump sites |
| 1.3.3 Construction sites |
| 1.4 Artificial non-agricultural vegetated areas | 1.4.1 Green urban areas |
| 1.4.2 Sport and leisure facilities |
| 1. Agricultural areas | 2.1 Arable lands | 2.1.1 Non-irrigated arable land |
| 2.1.2 Permanently irrigated land |
| 2.1.3 Rice fields |
| 2.2 Permanent crops | 2.2.1 Vineyards |
| 2.2.2 Fruit trees and berry plantation |
| 2.2.3 Olive groves |
| 2.3 Pastures | 2.3.1 Pastures |
| 2.4 Heterogeneous agricultural areas | 2.4.1 Annual crops associated with permanent crops |
| 2.4.2 Complex cultivation patterns |
| 2.4.3 Land principally occupied by agriculture with significant areas of natural vegetation |
| 3 Forests and semi-natural areas | 3.1 Forest | 3.1.1 Broad-leaved forest |
| 3.1.2 Coniferous forest |
| 3.1.3 Mixed forest |
| 3.2 Shrub and/or herbaceous vegetation association | 3.2.1 Natural grassland |
| 3.2.2 Moors and heathland |
| 3.2.3 Sclerophyllous vegetation |
| 3.2.4 Transitional woodland shrub |
| 3.3. Open spaces with little or no vegetation | 3.3.1 Beaches, dunes, and sand plains |
| 3.3.2 Bare rock |
| 3.3.3 Sparsely vegetated areas |
| 3.3.4 Burnt areas |
| 3.3.5 Glaciers and perpetual snow |
| 4 Wetlands | 4.1 Inland wetlands | 4.1.1 Inland marshes |
| 4.1.2 Peatbogs |
| 4.2. Coastal wetlands | 4.2.1 Salt marshes |
| 4.2.2 Salines |
| 4.2.3 Intertidal flats |
| 5. Water bodies | 5.1 Inland waters | 5.1.1. Water courses |
| 5.1.2. Water bodies |
| 5.2 Marine waters | 5.2.1 Coastal lagoons |
| 5.2.2 Estuaries |
| 5.2.3 Sea and ocean |

Table 27: CORINE land cover categories. Source: Copernicus Land Monitoring Service web page

# Annex II: Table of figures

[Figure 1: Transition between marginality and the definition on marginal lands in the framework of MAIL project. 9](#_Toc30586409)

[Figure 2: Different square areas where neighborhood analysis will be performed. In yellow is remarked the pixel that is computing the ratio. Source: personal compilation. 21](#_Toc30586410)

[Figure 3: The *MAIL* methodology overview for mapping MLs. 25](#_Toc30586411)

[Figure 4: Workflow of the methodology for mapping and classifying MLs in *MAIL*. Yellow parallelograms depict datasets used in the weighted overlay, while pink parallelograms depict “synthetic” indicators/layers. 26](#_Toc30586412)

[Figure 5: Overview of the indicator and criteria selection in the *MAIL* project. 27](#_Toc30586413)

[Figure 6: The elimination steps followed for selecting the final datasets used in *MAIL*. 49](#_Toc30586414)

[Figure 7: Datasets and layers as seen in the ArcGIS project. 55](#_Toc30586415)

[Figure 8: The final names of the Datasets v2 GIS layers. 55](#_Toc30586416)

Do not forget to update the Table of Figures

# Annex III: List of Tables

[Table 1: Factors for marginal land identification under *MAIL* scope. Source: Personal compilation 10](#_Toc30586417)

[Table 2: Constrains for marginal land identification under MAIL scope Source: Personal compilation 11](#_Toc30586418)

[Table 3: Technical specifications of CLC dataset. Source: Deliverable D2.2 12](#_Toc30586419)

[Table 4: Technical specifications of EUROSTAT dataset. Source: Deliverable D2.2 13](#_Toc30586420)

[Table 5: Technical specifications of Pan-European High Resolution Layers (HRL) dataset. Source: Deliverable D2.2 15](#_Toc30586421)

[Table 6: Technical specifications of Digital European elevation model dataset. Source: Deliverable D2.2 16](#_Toc30586422)

[Table 7: Technical specifications of Environmental dataset. Source: Deliverable D2.2 17](#_Toc30586423)

[Table 8: Technical specifications of the European Soil Database Derived data dataset. Source: Deliverable D2.2 18](#_Toc30586424)

[Table 9: Main specifications of the datasets used. Source: Personal compilation 19](#_Toc30586425)

[Table 10: The subset of 24 studies on which the indicator selection was based on. 27](#_Toc30586426)

[Table 11: Soil and terrain marginality indicators used in *MAIL*. 29](#_Toc30586427)

[Table 12: Sustainability marginality indicators used in *MAIL*. 29](#_Toc30586428)

[Table 13: Productivity marginality indicators used in *MAIL*. 29](#_Toc30586429)

[Table 14: Terrain and soil indicators and thresholds, sorted by importance (times used in literature review) 34](#_Toc30586430)

[Table 15 Sustainability indicators and thresholds, sorted by importance (times used in literature review) 37](#_Toc30586431)

[Table 16 Productivity indicators and thresholds, sorted by importance (times used in literature review) 40](#_Toc30586432)

[Table 17 Climate indicators and thresholds, sorted by importance (times used in literature review) 42](#_Toc30586433)

[Table 18: The ranks given to each indicator, in order to calculate the weights. 43](#_Toc30586434)

[Table 19: The pairwise comparison matrix of the marginality indicators, based on the ranks. 45](#_Toc30586435)

[Table 20: The normalized pairwise comparison matrix and the calculated weights. 46](#_Toc30586436)

[Table 21: Terrain and Soil indicator classes, scores and weights. 47](#_Toc30586437)

[Table 22: Sustainability indicator classes, scores and weights. 47](#_Toc30586438)

[Table 23: Productivity indicator classes, scores and weights. 48](#_Toc30586439)

[Table 24: Final datasets used in *MAIL*. 50](#_Toc30586440)

[Table 25: Additional datasets for use in *MAIL*. 52](#_Toc30586441)

[Table 26: Datasets for potential use in MAIL, as an addition to datasets from T2.2. 52](#_Toc30586442)

[Table 27: CORINE land cover categories. Source: Copernicus Land Monitoring Service web page 67](#_Toc30586443)

1. **R** = Report, **P** = Prototype, **D** = Demonstrator, **O** = Other [↑](#footnote-ref-1)
2. **PU** = Public, **PP** = Restricted to other programme participants (including the Commission Services), **RE** = Restricted to a group specified by the consortium (including the Commission Services), **CO** = Confidential, only for members of the consortium (including the Commission Services). [↑](#footnote-ref-2)
3. The EU has revised its legislative framework in order to meet the requirements for climate change mitigation under the 2015 Paris Agreement. As part of this framework the Land Use, Land Use Change and Forestry (LULUCF) regulation (2018/841) was adopted in May 2018 (European Parliament, 2018). [↑](#footnote-ref-3)
4. Land use is proposed as variable to detect other uses, as extensive grazing, that could compete with the uses proposed under the scope of MAIL project. Land uses classes will be analyzed regionally by specialists in order to detect uses that will constrain afforestation or other actions proposed by MAIL. [↑](#footnote-ref-4)
5. Variables related with extra factors (mainly soil variables) will be utilized in order to improve marginal land detection under the scope of MAIL project. [↑](#footnote-ref-5)
6. Minimum parcel size of marginal land is subject to review according the first outputs of the detection methodology and the knowledge depicted from pilot cases. [↑](#footnote-ref-6)
7. Regulation 2018/841 adopted in May 2018. This normative establish a land-based approach for accounting the emissions and removals from the LULUCF sector in five land accounting categories: (1) afforested and forested land; (2) managed cropland, grassland and wetland; (3) managed forest land; (4) harvested wood products; and (5) natural disturbances (Romppainen, 2019). [↑](#footnote-ref-7)
8. Area of STU allocation, Depth available to roots, Clay content (topsoil & subsoil), Sand content (topsoil & subsoil), Slit content (topsoil & subsoil), Organic carbon content (topsoil & subsoil), Bulk density (topsoil & subsoil), Coarse fragments (topsoil & subsoil), Total water content from PTR and PTF (topsoil & subsoil) [↑](#footnote-ref-8)
9. European Economic Area (EEA) 39: 33 member countries and six cooperating countries. The 33 member countries are the 28 European Union Member States, together with Iceland, Liechtenstein, Norway, Switzerland and Turkey. The six cooperating countries are Albania, Bosnia and Herzegovina, Kosovo\*, Montenegro, North Macedonia and Serbia. [↑](#footnote-ref-9)
10. Total available water content, Depth available to roots, Clay content, Silt content, Sand content, Organic carbon, Bulk Density, Coarse fragments [↑](#footnote-ref-10)
11. Detailed description concerning the calculation of the marginality index and the reclassification of the CLC classes required can be found in Annex I: Marginality index calculation and CLC reclassification. [↑](#footnote-ref-11)
12. Size of the square used for neighborhood analysis will be evaluated taking into consideration the spatial resolution of the land cover raster (10x10m for S2GLC dataset), and the scale where analysis is performed (local, national, regional or European). [↑](#footnote-ref-12)
13. Regulation 2018/841 adopted in May 2018. This normative establish a land-based approach for accounting the emissions and removals from the LULUCF sector in five land accounting categories: (1) afforested and forested land; (2) managed cropland, grassland and wetland; (3) managed forest land; (4) harvested wood products; and (5) natural disturbances (Romppainen, 2019). [↑](#footnote-ref-13)
14. <https://global-surface-water.appspot.com/> [↑](#footnote-ref-14)
15. <https://earthenginepartners.appspot.com/science-2013-global-forest> [↑](#footnote-ref-15)
16. <http://www.climatologylab.org/terraclimate.html> [↑](#footnote-ref-16)
17. <https://www.esri.com/en-us/arcgis/about-arcgis/overview> [↑](#footnote-ref-17)
18. <https://desktop.arcgis.com/en/arcmap/10.3/tools/data-management-toolbox/create-raster-catalog.htm> [↑](#footnote-ref-18)
19. <https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/creating-a-mosaic-dataset.htm> [↑](#footnote-ref-19)
20. <http://resources.arcgis.com/en/help/main/10.1/index.html#//01m10000000r000000> [↑](#footnote-ref-20)
21. Classes to be included into the final ratio will depend on the final definition of each land cover utilized by CLC project for the compilation of the European land use dataset. [↑](#footnote-ref-21)
22. For more information regarding the definition of each class, please consult “CORINE land cover nomenclature illustrated guide” [↑](#footnote-ref-22)