Developing support for monitoring and reporting of GHG emissions and removals from land use, land use change and forestry

Deliverable date: 30 April 2021

Submission date: January,2023

Edited by Ali Nadir Arslan (FMI)

Contributors:

SRTI-BAS: Kameliya Radeva, Lachezar Filchev

CASTRA: Vesselin Vasilev

CUNI: Premysl Stych

FMI: Tuula Aalto, Hannakaisa Lindqvist, Maria Tenkanen

SYKE: Iida Autio, Markus Törmä

LUKE (FMI's subcontractor): Tarja Tuomainen, Markus Haakana

MU: Stephen Barry

IGIK: Katarzyna Dabrowska-Zielinska, Agata Hoscilo

CBK-PAN: Stanislaw Lewinski, …

IHCANTABRIA: Jose Manuel Alvarez-Martinez

https://www.copernicus-user-uptake.eu/

Table of Contents

INTRODUCTION **3**

1. REVIEW ON EU INVENTORY SYSTEM QA/QC CURRENTLY EMPLOYED AT NATIONAL LEVEL BY MEMBER STATES **3**

1.1. BULGARIA 3
1.2. CZECH REPUBLIC 7
1.3. FINLAND 9
1.4. IRELAND 14
1.5. POLAND 17
1.6. SPAIN 19

2. EUROPEAN COPERNICUS DATA FOR NATIONAL LULUCF REPORTING **25**

3. PROPOSED SOLUTIONS AND METHODOLOGIES FOR IMPROVING THE CURRENT EXISTING LULUCF METHODOLOGY AND REGULATIONS **36**

3.1. USER ENGAGEMENT 36
3.2. TEST CASES IN SELECTED NUTS 40
3.3. PROPOSED SOLUTIONS AND METHODOLOGIES FOR USE IN NATIONAL REPORTING 41
3.3.1. Bulgaria (SRTI-BAS & CASTRA) 41
3.3.2. Czech Republic (CUNI) 47
3.3.3. Finland (FMI (Luke & INRAE) & SYKE) 50
3.3.4. Ireland (MU) 65
3.3.5. Poland (IGIK & CBK-PAN) 71
3.3.6. Spain (IHCANTABRIA) 81

4. IMPLICATIONS OF COPERNICUS PRODUCTS FOR REPORTING **177**

5. REFERENCES **1**

Introduction

This report is final report from the Action 2019-2-49 of the Caroline Herschel Framework Partnership Agreement (FPCUP), entitled," Developing support for monitoring and reporting of GHG emissions and removals from land use, land use change and forestry".

1. Review on EU inventory system QA/QC currently employed at national level by Member States

According to the 2006 IPCC GL the QA/QC system, that should be implemented for GHG Inventories consists of an inventory agency responsible for coordinating QA/QC activities, a QA/QC plan, general QC procedures (Tier 1), source category-specific QC procedures (Tier 2), QA review procedures and verifications as well as procedures regarding reporting, documentation, and archiving.

The QA/QC plan is a basic element of the QA/QC system. The plan outlines QA/QC activities that are implemented and includes the scheduled time frame for inventory preparation from its initial development through the final reporting in any year. It contains an outline of the processes and schedule to review of all source categories.

The QA/QC plan is an internal document to organise, plan and implement QA/QC activities. Once developed for the next submission, it is referenced and used in subsequent inventory preparation, or modified as appropriate.

The QA/QC plan is provided for implementation to all institutions, which are engaged in the process of preparation of emissions inventories under UNFCCC as provision of the relevant activity data.

Based on the national QA/QC plan each of the institutions has nominated experts, responsible for preparation of the required information as well as for implementation of QA/QC procedures.

1.1. Bulgaria

As illustrated in *Figure 1.1* and outlined shortly the Bulgaria's reporting obligations to the UNFCCC, UNECE and EC are being administered by the Ministry of environment and water (MoEW). All activities on preparation of GHG inventory in Bulgaria are coordinated and managed on the state level by MoEW. The Executive environment agency (ExEA) , that is a structure of MoEW, has been identified as the responsible organization for preparation of Bulgaria's National GHG Inventory under the UNFCCC and the Kyoto Protocol and designated as single national entity. ExEA has the technical responsibility for the national inventory: acts as National Inventory Compiler (supervises inventory preparation process); manages Bulgarian National Inventory System (BGNIS); compiles CRF tables and national inventory reports (NIR); coordinates the work of engaged consultants for supporting inventory; coordinates and implements the activity of National QA/QC Plan.

The Bulgarian Quality Management System was established in 2008. The QA/QC plan is an internal document to organise, plan and implement QA/QC activities. Once developed for the next submission, it is referenced and used in subsequent inventory preparation, or modified as appropriate. The QA/QC plan has been updated in 2014 to implement the new established legal, institutional, and procedural arrangements within the Bulgarian

National Inventory System. The updated National QA/QC Plan was approved by the Ministry of Environment and Water in December 2014.

National QA/QC Plan includes elements like Responsible institutions;Data collection;Preparation of inventory;Category-specific QC procedures;QA and review procedures;Uncertainty analyses;Organisation of the activities in quality management system;Verification activities;Reporting, documentation and archiving.

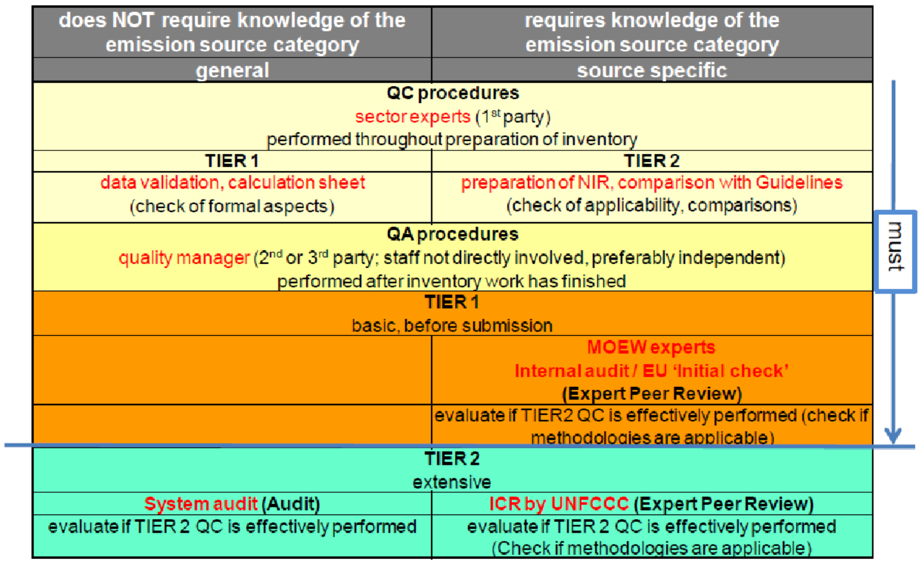


Figure 1.1: QA / QC procedures (source: ExEA_GHG emission inventory report 2019)

The QC procedures are performed by the sectors, that are directly involved in the process of preparation of inventory with their specific responsibilities. The quality assurance process includes expert review was conducted in two stages: a review of the initial set of emission estimates and, a review of the estimates and text of the Inventory Report. QA experts could be:

- Sector experts from the MoEW, which are engaged through internal administrative order by the MoEW;
- Experts from research institutes in accordance with their competence;
- Other external reviewer (national and/or international).

QA/QC procedures for LULUCF sector

The input data, estimates and results are checked as follows. • Bottom-up check • Input data • Check for the plausibility of the activity data and their trend • Check for plausibility of the emission factors as well as the related input data and their trends • Check of input data for completeness • Estimations • Check of the correctness of all equations in the estimate files • Check of the correctness of all interim results • Check of the plausibility of the results and their trends • Check of the correctness of all data and results transfer • Top-down check • Check of the consistence of the total area for Bulgaria.

A comparison of the activity data used with those from other statistics and a Comparison of the used emission factors and underlying input data with those of other data sources (e.g. from literature, results in NIRs of other comparable regions, IPCC default values) is also included in the process. In terms of QA/QC of the activity data, the correctness of the data on the areas and the tree stock is controlled during the preparation, the adoption and the execution of the Forest Management Plans (FMP). The quality control is exercised by the Executive Forest Agency and its subdivisions. The Quality control could be exercised by other institutions like Ministry of Environment and Waters, municipal authorities as well as by forest landowners. Quality control is exercised at every phase of the preparation of the FMP and the results of the check are documented and the mistakes are corrected. Concerning the agrostatistical data, from the Agrostatistics (BANSIK) and Strategies Directorate of MAFF together with the Regional Directorates "Agriculture and forestry" and Municipal Services on agriculture and forestry at MAFF organized and conducted the agricultural census in Bulgaria. Around 4000 surveyors participated in the data collection process. Around 400 controllers supervised the work of the surveyors and provided methodological assistance. The controllers delivered the checked questionnaires to the agrostatistics experts from the Regional Directorates "Agriculture and Forestry" according to a previously adopted schedule. The operators do the data entry in the census software spread in the regional offices. The regional data bases are aggregated on national level by Agrostatistics and Strategies Directorate of MAFF. The data entry from the filled in questionnaires into computer software was followed by crosschecks and coherence control in order to ensure the data quality.

Bulgaria uses a combination of different datasets available in the country for reporting of LULUCF. Bulgaria applies Approach 1 and 2 in land representation. Land use matrices are developed based on information on total area for each individual land use category and partial information on the net losses and net gains of each land use category. The information used for GHGI does not have the locations of the land use and land use changes.

The following datasets are available in Bulgaria with the potential of increasing the accuracy and completeness of the land representation system towards a geo-referenced approach:

- BANCIK – Land cover/Land use survey in Bulgaria since 1998. Annual evaluation based on two detailed nomenclatures – physical nomenclature (providing information on the land cover and vegetation of the observed point and functional nomenclature providing information on the land use (socio-economic dimension of the observed territory. The territory of the country is divided on 3123 segments distanced at 6 km. Each segment has 36 points spaced 234 m apart.
- IACS-LPIS - The Agricultural Land Parcel Identification System (LPIS) is a part of the Integrated Administration and Control System (IACS), which has been developed in all Member States following the main EU regulations. The LPIS in Bulgaria is developed based on a digital orthophoto map of aerial/satellite photography. The reference plot is a physical block. LPIS information is available since 2007 covering in a "wall to wall" the entire territory of Bulgaria. In case of LPIS, one-fourth of the country is systematically updated every year. From 2020 onwards, it is expected that one-third of the country will be updated each year. The main benefits of the dataset are the accurate spatial representation (1:5000 scale), the explicit management of temporal information and the thematic accuracy (classification correctness) of agricultural area (managed cropland and managed grassland).
- Map of the restored property ownership (since 1995) – combines the data from the land division plan, the map of the existing old real borders, the map of the restored old real borders of agricultural lands created under the Law on the Property and Use of Agricultural Land (LAPP), and the map of the restored property on Forests and lands from the forest fund, established by the order of the Law for restoration of ownership of forests and lands from the forest fund
- Executive Forest Agency – owns digitalized information on forest area as resulting from the Forest Management Plans and that contains information on (i) Forested areas. (ii) Non-forest areas for

afforestation, which represent forest territory which is temporary unstocked due to fire disturbances or logging activities, (iii) non-forested areas – forest roads, nurseries, meadows and pastures, croplands, rocks.

- National Statistic Institute – owns information on Cropland and Grassland areas for the years before 1998. The information is not georeferenced, and it is stored in the National Statistical Yearbooks.

The use of Copernicus data in current system of LULUCF reporting

The EU LULUCF Regulation (Regulation (EU) 2018/841) envisages the use of geographically explicit approaches to derive the land use and land-use changes for compilation of the GHGI in LULUCF. Currently Bulgaria applies a combination of Regulation's defined Approach 1 and Approach 2 in elaboration of land-use change matrices. Although geo-spatial data is available in the country many technical and administrative challenges burden the use of these resources. Moving towards the defined in the regulation Approach 3 for land use representation would increase the accuracy of the estimates in terms of carbon emissions and removals. Copernicus data has not been used in the GHG emission inventory process.

For verification of data assumptions regarding the land use changes and to double check some of the information regarding specific classes of land use which could be classified differently into the varied information systems in the country, only information from Corine Land Cover is used.

This implies the following main requirements for use of geospatial information and specific needs for LULUCF in Bulgaria:

Land representation under LULUCF:

1. Need for a complete and consistent land representation over time – it requires availability and access to historical data on land use and land use changes back to 1968 (for Bulgaria under UNFCCC)
2. Segmentation of minimum mapping units that meets the definition of forest in Bulgaria – (0.1 ha)
3. Need to differentiate managed versus non-managed lands – relevant for CL and GL in Bulgaria

Existing gaps in the national reports are the lack of complete and accurate data set and time series consistency. Very often there is lack of synchronization between the different information systems which is their main disadvantage. Main challenges here are related to:

- Lack of systematically collected information on land use changes between different land use classes during the years.
- Discontinuity of some statistics.
- Use of different definitions and terms
    - For example, some institutions store and maintain the information based on the designation of the land parcels whereas others work with the actual land use. All these specifics are considered in the process of land representation to ensure the consistency in definitions and land use classes. The activity data was threatened in hierarchical order when LUC matrices were elaborated to ensure the accuracy as much as possible.
- Activity data on other non-agricultural lands

- As these data sources store information for the entire country territory and not only on agricultural lands. They are also used to present information on WL, SM and OL.

- Additional data on land cover
    - Data from Corine Land Cover is used to verify the assumptions regarding the land use changes and to double check some of the information regarding specific classes of land use which could be classified differently into the varied information systems in the country.

The following main current challenges implementing geographically explicit methods in elaboration of land use matrices in Bulgaria has been identified by the ExEA generalised in two groups

First group: Resources and expertise
- Lack of human and financial resources in the institutions responsible for reporting of GHGI
- Lack of expertise within the LULUCF team for processing of remote sensing and geo-spatial information for development of land-use change matrices

Second group: Technical challenges
- Different temporal coverage of the spatially explicit data available in the country.
- Difficulties to construct a complete and consistent time series
- Different spatially explicit datasets challenge the processing, harmonization, and validation of the data.

Stakeholders –the Executive Environment Agency; Ministry of Environment and Water; Ministry of Agriculture, Food and Forestry; Executive Forest Agency, National Statistics Institute; Bulgarian Academy of Sciences; Universities, the Private sector – SMEs and professional NGOs working in the domain of EO data processing or ICT product development for Agriculture or other applications; World bank and other financing institutions, and consulting companies implementing infrastructural projects in which GHG/LULUCF emissions estimates are required as function of projects' execution, and other.

The FPCUP project will contribute to finding solutions to improve the existing methods used in national inventory GHG emissions reports like providing information for the sizes of main categories (forest land, cropland, grassland, wetlands, settlements) areas for LULUCF sector. Enhancing the use of Copernicus products and services will help to establish changes from one to another category or subcategory and overcome the lack of systematically collected information on land use changes between different land use classes during the years. The FPCUP project will boost the process for providing information to overcome the local need of a complete and consistent land representation over time and the need to differentiate managed versus non-managed land by using Copernicus data, starting from a certain suitable year for Bulgarian case.

1.2. Czech Republic

The methodology requirements and principles associated with the approaches recommended by the GPG for LULUCF (IPCC 2006 Gl. (IPCC 2006)) imply that, for the reported period from 1990, the required land use should be available for the period starting from 1969. Consistent representation of land areas and identification of land-use changes constitute the key steps in the inventory of the LULUCF sector in accordance with the IPCC 2006 Gl.

(IPCC 2006). The adopted system of land-use representation and land-use change identification was constructed gradually. Since the 2008 NIR submission, this has been exclusively based on the cadastral land use information of the Czech Office for Surveying, Mapping and Cadastre (COSMC; www.cuzk.cz). The Czech land-use representation and the land-use change identification system use annually updated COSMC data, elaborated at the level of about 13 thousand individual cadastral units. The system was constructed in several steps, including 1) source data assembly 2) linking land-use definitions 3) identification of land-use change 4) complementing time series. The result is a system of consistent representation of land areas having the attributes of both Approach 2 and Approach 3 (IPCC 2006), permitting accounting for all land-use transitions in the annual time step. Information of land use administers the database of "Aggregate areas of cadastral land categories" (AACLC). The AACLC data are compiled at the level of the individual cadastral units. There are over 13 000 cadastral units, the number of which varies due to separation or division for various administrative reasons. COSMC provides the annually updated areas for all land-use categories, the Forest Management Institute (FMI) reports the recent data on forests (harvest, increment, felling, etc.) that are used in the land-use categories involving forest land. The preparatory calculation is mostly performed in excel spreadsheets. The final data files include the checked and verified data. The current submission covers the whole reporting period from the base year from 1990. For every year since 2000 the authorized institutions in Czechia create NATIONAL GREENHOUSE GAS INVENTORY REPORT OF THE CZECH REPUBLIC and they are available from here: http://portal.chmi.cz/files/portal/docs/uoco/oez/nis/nis_do_cz.html. The results for whole period of reporting consists of Excel sheets with data for each category of observation like Energy, LULUCF, Agriculture, Waste etc. These results are also available from this website http://portal.chmi.cz/files/portal/docs/uoco/oez/nis/nis_do_cz.html.

The use of Copernicus data in current system of LULUCF reporting

Czech Republic uses cadastral data from Czech Office for Surveying, Mapping and Cadastre (COSMC) for reporting of LULUCF. Copernicus data is not used in the current system of LULUCF reporting.

Existing gaps in the national reports are no Earth observation implemented, data based on Czech Office for Surveying, Mapping and Cadastre (COSMC) are applied. This data has a limited capacity to detect dynamic changes. Moreover the archive data miss the spatial information registered in GIS. Land use nomenclature/legend of Czech Office for Surveying, Mapping and Cadastre is limited (10 classes) and some classes combine together two or more LULUCF categories, e.g. class "other" in COSMC includes areas with settlements and other based on LULUCF.

Stakeholders for national reporting data are Ministry of the Environment of the Czech Republic. The Czech Hydrometeorological Institute (CHMI), under the supervision of the Ministry of the Environment, is designated as the coordinating and managing organization responsible for the compilation of the national GHG inventory and reporting its results. Institute of Forest Ecosystem Research Ltd. (IFER), Jilove u Prahy, is responsible for compilation of the inventory in the sector of Land Use, Land Use Change and Forestry. The procedure of inventory compiling is initiated by IFER. IFER collects the required data from the Czech Statistical Office (CzSO), the Czech Office for Surveying, Mapping and Cadastre (COSMC) and the Forest Management Institute (FMI).

1.3. Finland

The annual inventory and reporting of greenhouse gas emissions and removals provide an information base for the planning and monitoring of climate policy. The Kyoto Protocol obliges its parties to establish a national greenhouse gas inventory system by the end of 2006. Finland's National Greenhouse Gas Inventory System was set up at the beginning of 2005 (Official Statistics of Finland (OSF) 2021) .

Finland has ratified and is committed to follow the United Nations Framework Convention on Climate Change (UNFCCC), the Kyoto Protocol and its Doha Amendment as well the Paris Agreement. In addition, Finland as an EU Member State, has obligations related to climate change under EU legislation.

Under the EU, UNFCCC and the Kyoto Protocol, Finland is required to submit annually a national greenhouse gas inventory (GHGI) covering emissions and removals of direct greenhouse gases from the five sectors (Energy, Industrial Processes and Product Use, Agriculture, Land Use, Land-Use Change and Forestry and Waste) and for all years from the base year to the most recent year. The preparation and reporting of the Finnish GHGI are guided by the UNFCCC reporting guidelines (UNFCCC 2013) implementing the methodological guidance in the 2006 IPCC Guidelines for National Greenhouse Gas inventories. Finland has not elected the KP LULUCF activity *Wetland Drainage and Rewetting.* Therefore, Finland has not used the 2013 Supplement to the 2006 IPCC guidelines for National Greenhouse Gas Inventories: Wetlands in the inventory preparation except in a few cases, where the IPCC Wetlands Supplement has been used as a reference when updating national emission factors for drained organic soils in both the Agriculture and LULUCF sectors (Official Statistics of Finland (OSF) 2021).

According to the Government resolution of 30 January 2003 on the organisation of climate policy activities of Government authorities, Statistics Finland assumed the responsibilities of the national entity for Finland's greenhouse gas inventory from the beginning of 2005. In 2015, the role of Statistics Finland as the national entity was enforced through the adoption of the Climate Change Act (609/2015).

In Finland, the national system is established on a permanent footing and it guides the development of emission calculation in the manner required by the Kyoto Protocol. The national system is based on laws and regulations concerning Statistics Finland, on agreements between the inventory unit and expert organisations on the production of emission and removal estimates, as well as related documentation. Statistics Finland has also agreements on cooperation and support to the expert organisations participating in Finland's national system with relevant ministries. The national system is designed and operated to ensure the transparency, consistency, comparability, completeness, accuracy and timeliness of greenhouse gas emission inventories. The quality requirements are fulfilled by implementing consistently the inventory quality management procedures. The national system for the greenhouse gas inventory in Finland is presented in *Figure 1.2* (Official Statistics of Finland (OSF) 2021).

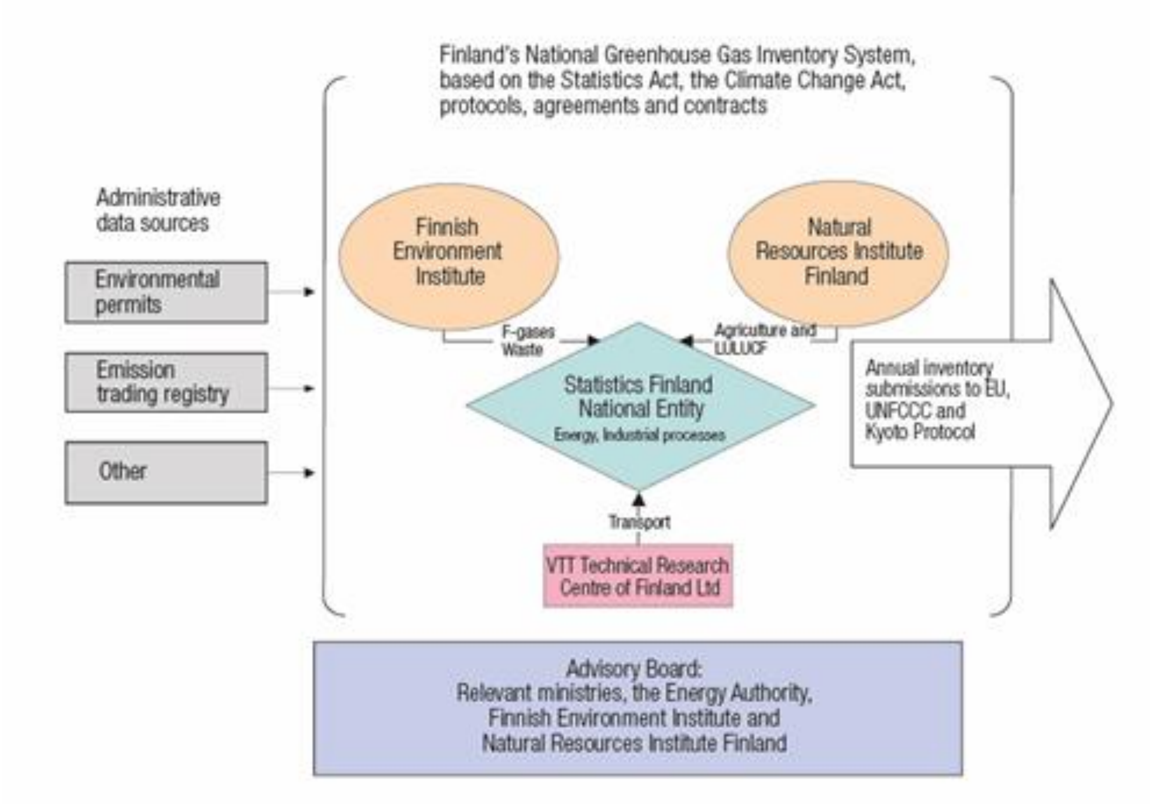


Figure 1.2: The National System for the Greenhouse Gas Inventory in Finland (LULUCF = Land use, land-use change and forestry). Natural Resources Institute Finland (Luke) is responsible expert organisation for reporting LULUCF sector (Official Statistics of Finland (OSF) 2021).

Quality management (incl. quality objectives and QA/QC and verification plan in national level), sector specific QA/QC details are discussed in the relevant sections of National inventory Report (NIR) (Official Statistics of Finland (OSF) 2021).

The inventory process consists of four main stages: planning, preparation, evaluation and improvement (PDCA cycle) and aims at continuous improvement. A clear set of documents is produced on the different work phases of the inventory. The documentation ensures the transparency of the inventory: it enables external evaluation of the inventory and, where necessary, its replication (*Figure 1.3*).

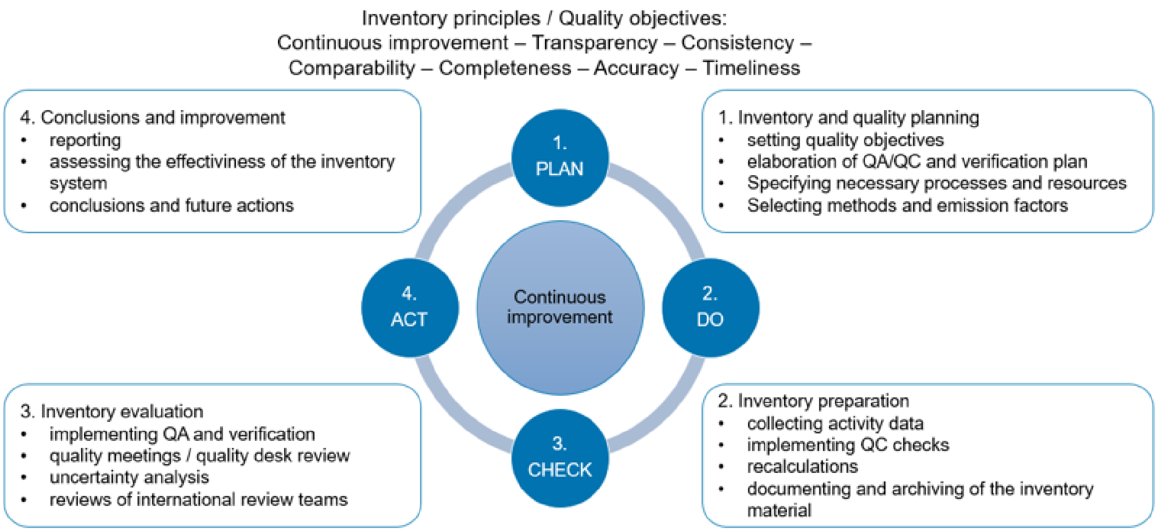


Figure 1.3: Inventory and QA/QC process of Finnish GHG Inventory (Official Statistics of Finland (OSF) 2021).

Quality objectives are specified statements about the quality level that is aimed at the inventory preparation about the inventory principles. The objectives aim to be appropriate and realistic while considering the available resources and other conditions in the operating environment.

Quality objectives for all sectors in the Finnish GHG Inventory (Official Statistics of Finland (OSF) 2021) are given below

1. *Continuous improvement*
    1.1. Treatment of review feedback is systematic
    1.2. Improvements promised in the National Inventory Report (NIR) are carried out
    1.3. Improvement of the inventory is systematic
    1.4. Inventory quality control (QC) procedures meet the requirements
    1.5. Inventory quality assurance (QA) is appropriate and sufficient
    1.6. Verification of the inventory meet the requirements
    1.7. Known uncertainties of the inventory are taken into consideration when planning the improvement needs
2. *Transparency*
    2.1. Archiving of the inventory is systematic and complete
    2.2. Internal documentation of calculations supports emission and removal estimates
    2.3. CRF tables and the National Inventory Report (NIR) include transparent and appropriate descriptions of emission and removal estimates and of their preparation
3. *Consistency*
    3.1. The time series are consistent
    3.2. Data have been used in a consistent manner in the inventory
4. *Comparability*
    4.1. The methodologies and formats used in the inventory meet comparability requirements
5. *Completeness*
    5.1. The inventory covers all the emission sources, sinks, gases and geographic areas

6. *Accuracy*
    6.1. Estimates are systematically neither higher nor lower than the true emissions or removals
    6.2. Calculation is correct
    6.3. Inventory uncertainties are estimated
7. *Timeliness*
    7.1. High-quality inventory reports reach their receivers (EU/UNFCCC) within the set time

Internal and external audits

An annual in-depth-review of the inventory by sector or responsibility area is done mainly in conjunction with the bilateral quality meetings or the quality desk review. The bilateral quality meetings are held annually between the inventory unit (the compiler) and the expert organisations (producing the inventory estimates and descriptions) in January to February. In 2021 (inventory 2019) bilateral quality meetings were held for every sector.

The main objective of the quality meetings and quality desk review is to ensure that the experts have implemented the QC checks and required QA and verification procedures according to the QA/QC and verification plan and to evaluate the results and documentation of the procedures (*Figure 1.4*). Quality meetings and desk reviews follow a fixed agenda that include the following items: Implementation of the QA/QC plan, category-specific QA/QC and verification actions if relevant, review feedback, structure and transparency of the reporting (NIR and CRF tables), improvement needs and plans, and functioning of the national inventory system (e.g. resources for inventory preparation) (Official Statistics of Finland (OSF) 2021).

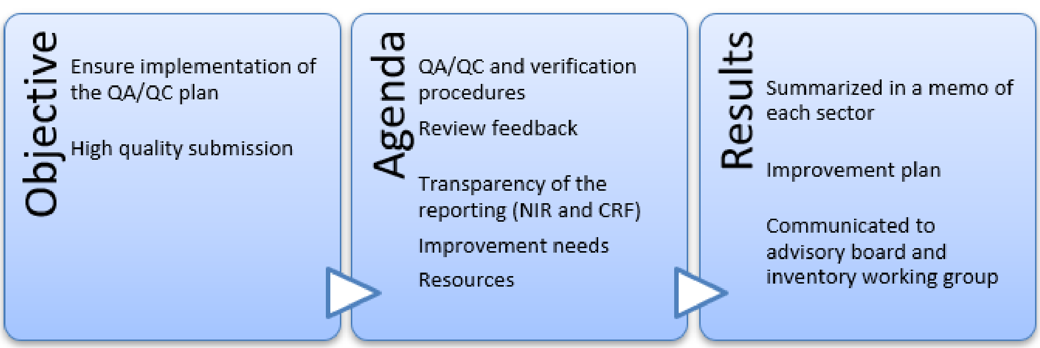


Figure 1.4: Objective, agenda and results of the national quality meetings and desk reviews.

Finland's inventory system has a special procedure for the consideration and approval of the recalculations (*Figure 1.5*). If sectoral experts identify any needs for recalculations, they contact the inventory unit and provide comparison calculations and solid justification for the recalculation. The methodological changes are then communicated to the advisory board for evaluation and approved by the inventory unit before adopted into production.

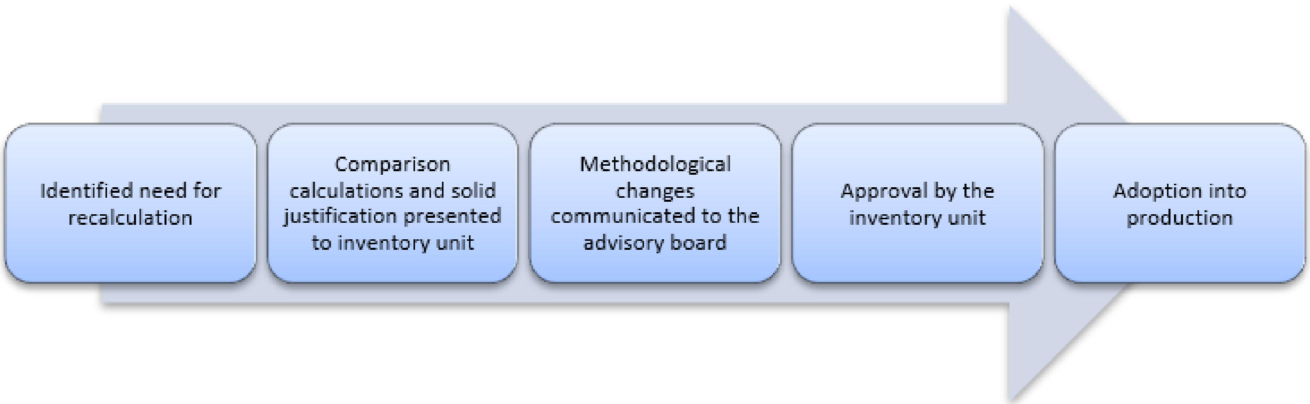


Figure 1.5: Procedure for consideration and approval of the recalculations.

A management team of experts from Luke supervise the reporting of emissions and removals for the LULUCF sector. The members have a broad expertise in using the measurements and methodology to estimate carbon stock changes and greenhouse gases. All changes in methods, activity data and emissions or other factors and parameters are discussed and approved by the management team before they are introduced to the advisory board. The management team meets two to four times per year. Category specific QA/QC plans are presented in the NIR (Official Statistics of Finland (OSF) 2021).

The methods and data used for LULUCF reporting can be divided into two separate processes:

1) Monitoring of LU categories and LU changes, methodology to estimate areas of LU and LUC, and how time series of LU and LUC are constructed.

2) Methodology to estimate emissions by sources and removals by sinks for each LU and LUC types separately.

The monitoring and estimation of LU and LUC areas are based on the national forest inventory (NFI) field sampling design. The sampling design enables the monitoring of all land use types, because it is not limited only to forest land, but sampling plots are assessed also on other land uses. The NFI data on soil type and land use changes are complemented with other spatial data like satellite images, aerial photographs, numerical maps, Finnish soil database, EU's Land Parcel Identification System and Building and Dwelling Register (BDR) of Digital and Population Data Services Agency (information on buildings, construction projects, residences, real estates).

Copernicus data is not used in the current GHGI. However, national high resolution spatial data is used in evaluation land-use change observations of NFI and to update land use recordings in NFI due to 5 years' inventory cycle of field measurements. Corine land cover layers have been used in research projects related to GHGI to examine potential of Copernicus products on land use monitoring.

Existing gaps are the unavailability of up-to-date data on land-use changes for the most recent years, and lack of data on tree biomass and its changes on other LUs than forest land. There is hardly no data for verification of emissions and removals on different LUs and/or gases, or for the whole LULUCF sector.

Stakeholders. Statistics Finland is the national entity with the overall responsibility for the compilation and finalisation of inventory reports and their submission to the UNFCCC Secretariat and the European Commission. Natural Resources Institute Finland (Luke), Finnish Environment Institute (SYKE), VTT Technical Research Centre of Finland Ltd, The Energy Authority, the Ministry of Economic Affairs and Employment and the Ministry for Foreign Affairs of Finland have responsibilities to prepare Finnish GHG Inventory. Ministry of the Environment and Ministry of Agriculture and Forestry are closely related to the Inventory. Current stakeholders are also National Land Survey of Finland, Finnish Food Authority, Centre for Economic Development, Transport and the Environment (North Ostrobothnia) as data providers to LULUCF reporting and Finnish Meteorological Institute (FMI) with other research organisations.

1.4. Ireland

In early 2005, the inventory agency in Ireland commissioned a project with UK consultants NETCEN to establish formal QA/QC procedures that would meet the needs of the UNFCCC reporting requirements. The project developed a QA/QC system including a documented QA/QC plan and procedures along with a QA/QC manual.

The QA/QC plan identifies the specific data quality objectives related to the principles of transparency, consistency, completeness, comparability, and accuracy required for Ireland's national inventory and provides specific guidance and documentation forms and templates for the practical implementation of QA/QC procedures. The QA/QC procedures cover such elements as data selection and acquisition, data processing and reporting.

The inventory agency initiated a new approach to QA/QC in the 2006 reporting cycle. Its application was completed and consolidated in delivering the submissions up to this present 2020 submission. This involved the allocation of responsibilities linked to the national system and the use of a template spreadsheet system to record the establishment and maintenance of general inventory checking and management activities covering the overall compilation process, as well as the undertaking of specific annual activities and any necessary periodic activities in response to specific events or outcomes in inventory reporting and review. The system facilitates record keeping related to the chain of activities from data capture, through emissions calculations and checking, to archiving and the identification of improvements.

The Irish inventory system's quality assurance and quality control process was reviewed by external contractors in 2012 and 2017. These reviews were considered an important element of the quality assurance and control process by the inventory agency and improved the transparency and consistency between Convention and KP reporting for LULUCF.

Additional LULUCF QA/QC procedures are implemented on the Forest Lands (4.A) sector. This process includes

- Evaluation of required data from external sources (Forest service, Collite) and the establishment of memoranda of understanding between the Department of Agriculture, Forestry and the Marine, the Environmental Protection Agency (the inventory agency) and data providers including:

  - Deadlines for data delivery;
  - Internalised QA/QC checks and procedures;
  - Metadata;
  - Notification of changes to methods used for collecting activity data;
  - Identification of contact points and responsible parties.
- Correspondence with data providers 2 months before agreed delivery dates to notify of new requirements, request notification of changes to any activity data and to remind providers of deadlines;
- QC checks of reference sources for national activity data by evaluation of documentation with regard to activity data. For example, is data collection or sampling regimes adequate and unbias? Does the agency have any information on uncertainties?
- Comparisons of input data with independent data sets such as harvest statistics (FAO/Eurostat), land cover data such as CORINE (see Black et al., 2009a);
- Time series consistency checks of activity data;
- Collation and initial completeness checks of activity data required;
- Pre-processing activity data and compiling data bases to be used by CBM (Forestry Carbon Model)

Emission Factors, Models and Calculations

QA/QC checks were performed on the background data used to develop emission factors. The Carbon Budget Model (CBM) and the FORCARB were used in preparation of emission inventories and are both widely used.

Both the FORCARB and CBM models were developed specifically for GHG inventory reporting. When these models were designed and developed the following was considered;

- Appropriateness of model assumptions, extrapolations, interpolations;
- Model calibration: models have been calibrated using National Forest Inventory data;
- Calibration of the age class distributions used in the FORCARB model was checked against independently derived information;
- Models are re-evaluated and updated annually using any new research or based on the output from uncertainty analysis
- All pools are included in the models, so are complete in relation to the IPCC source/sink categories.
- The calibrated CBM model has been cross-compared with the previous CARBWARE mode and is shown to improve CSC estimates for Irish forestry
- CBM has been validated against real time eddy covariance data and show good agreement in net ecosystem change estimates

Validation of CBM

The CBM was compared to real time Net Ecosystem Exchange (NEE) measurements and good agreement (r^2= 0.84) between CO_2 exchange from the CBM and NEE was observed (see Figure 1.6).

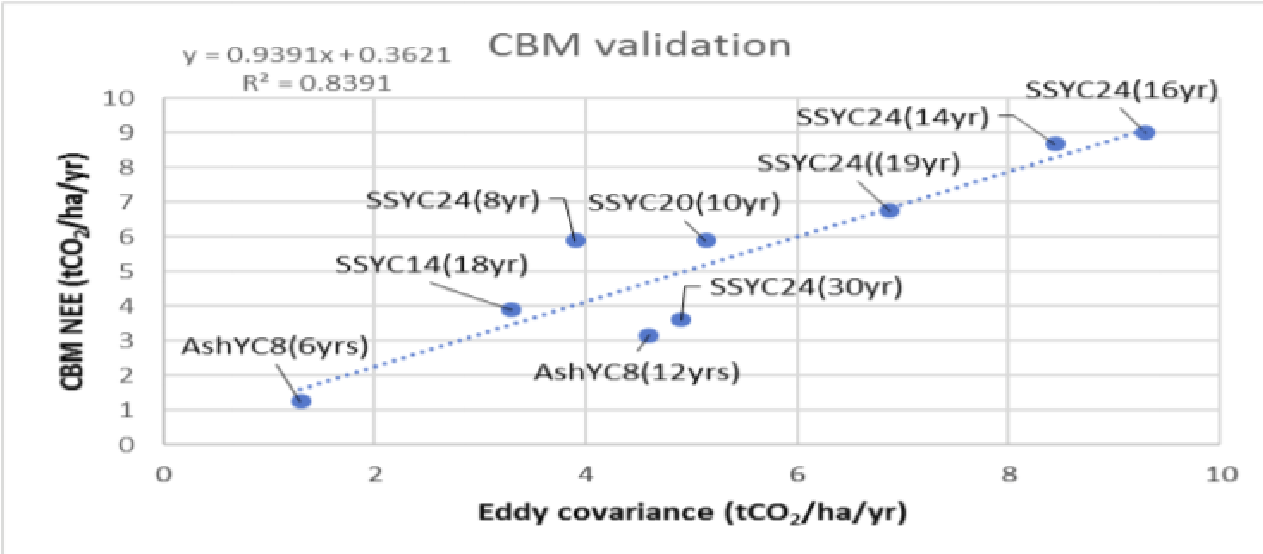


Figure 1.6: CMB validation against real time Net Ecosystem Exchange (NEE) measurements.

The CBM also provides estimates of biomass and DOM pools which are expected to agree with NEE estimates if the correct calibration parameters are used. The performance of the CBM estimate of biomass C stocks was assessed using various statistical

These measures included:

- Adjusted r2 and P value
- Percentage bias (Bias %)
- Root mean square error (RMSE)
- Percentage error at the 95% confidence interval (Error %)
- Students T test P value, where P values < 0.05 suggest that CBM and validation data are significantly different

Figure 1.7 shows that the CBM adequately represented the biomass of Validation of total biomass estimates from CBM and research data plots for different aged Sitka spruce, mixed Sitka spruce and Lodgepole pine (SS/LP)) and Ash stands.

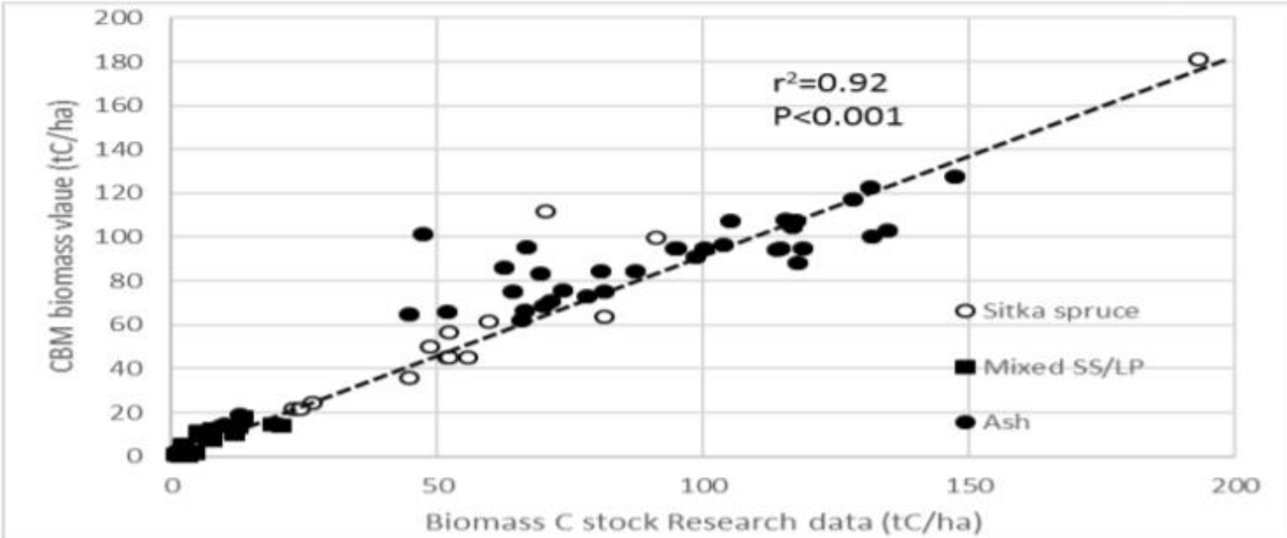


Figure 1.7: Validation of biomass between CBM and research data.

Ireland uses data from national road building statistics, and the area covered by new residential, commercial and industrial construction based on CSO and Department of Housing, Planning and Local Government annual construction statistics, which also report floor area of development projects. With the exception of Forest converted to Settlement, the identification of previous land use from which settlement areas are converted is based on an analysis of the distribution of land use classes given by CORINE 1990.

Existing gaps in the national reports include an implicit assumption of no carbon stock change in lands adjacent to new constructions (green areas, etc.) relative to previous land use. Using Copernicus products such as Sentinel 2 imagery, an assessment of lands adjacent to new constructions could provide information regarding land cover changes and using biophysical parameters it may be possible to identify changes in vegetation health and abundance (NDVI and LAI) to assess the impact of the construction on adjacent green areas.

Stakeholders for national reporting data may include the Central Statistics Office and d Department of Housing, Planning and Local Government who compile estimates.

1.5. Poland

Most of the input data used in the inventory process comes from official national statistics in the statistical studies of Statistics Poland, reports of Forest Management and Geodesy Bureau.

The QA/QC programme has been elaborated in line with the 2006 IPCC Guidelines for National GHG Inventories to ensure high quality of the Polish annual greenhouse gas inventory. The QA/QC programme contains tasks, responsibilities as well as time schedule for performance of the QA/QC procedures. The following elements of the Quality Assurance and Quality Control system are addressed:

- Inventory agency responsible for coordinating QA/QC activities,
- QA/QC plan,
- General QC procedures (Tier 1 method),
- Source category-specific QC procedures (Tier 2),
- QA review procedures,
- Reporting, documentation and archiving procedures.

QC covers routine technical activities carried out with the aim of quality control of national emissions and removals inventories allowing for: a) Maintaining the correctness and completeness of data, b) Elimination of errors and determination of potential deficiencies. QC activities contains checks for accuracy of data and estimations acquiring as well as application of approved procedures for calculation of emissions, uncertainty, archiving of information and reporting.

Activities aiming at quality assurance (QA) cover procedural systems for control carried out by experts not involved directly in elaborating GHG inventory in given sector. QA activities are conducted over a completed inventory and allow to ensure that national inventory represents top level of emissions and removals assessment at the present knowledge and available data and effectively support quality control (QC).

The basic elements of QA/QC plan are implemented and co-ordinated by the National Centre for Emission and Management (KOBiZE), the unit responsible for Polish GHG inventory preparation. It follows the 2006 IPCC Guidelines for National GHG Inventories recommendations. The main procedures for QA/QC activities are described in the National Quality Assurance / Quality Control and Verification Programme of the Polish Greenhouse Gas Inventory.

Depending on methodology used for emission estimation within categories Tier 1 or Tier 2 check procedures are carried out. The extended QC procedure for checking the correctness of emissions estimations is used for these categories where country specific emission factors are established. This concerns the key categories especially for such sectors like fuel combustion, transport, cement production, enteric fermentation, manure management, and others. For GHG emission sources for which Tier 1 method is used for emission calculation also Tier 1 method is applied for inventory checks (Poland's National Inventory Report 2020).

The use of Copernicus data in current system of LULUCF reporting

To date, the Corine Land Cover database is only used in the LULUCF reporting in Poland. The CLC is used for assessment of the wetland areas and as an additional dataset for delineation of agricultural land. The Copernicus High Resolution Layers has not been used yet in the inventory process.

Poland uses data from Head Office of Geodesy and Cartography and General Statistical Office for reporting of LULUCF.

Existing gaps in the national reports are as follow,

- Lack of information on LULUC on frequent bases from National Land Survey of Poland (GUGIK) which supply the cadastre data. There is the limitation on dynamic changes.
- There is the lack of biomass data and biophysical parameters which could be supplied from Copernicus data.
- The verification data does not exist. This could be supplied by EO data.
- The only information for the reporting which comes from Corine is information about wetlands.
- Wetlands are dynamic ecosystems and EO data could be used for discrimination of water which cover the grasslands temporarily.
- The most significant weakness of reference dataset related to forest land is that it implements an exclusive forest definition, which is 'land administration oriented' thus focusing on forest administrated by the State Forests Holding, instead of one based on quantitative thresholds which would be able to capture all forests and change no matter of their cadastral status.

Stakeholders for national reporting data are: The National Centre for Emissions Management (KOBIZE), Ministry of Environment, National Land Survey of Poland (GUGiK) and National Statistics (GUS).

IGIK has permanent contact with KOBIZE and GUS. IGIK organises workshops and meetings.

1.6. Spain

The Spanish System of Inventory and Projections of Emissions to the Atmosphere of Greenhouse Gases and Air Pollutants is elaborated by the Ministry for the Ecological Transition and the Demographic Challenge (MITERD), the department of the Government of Spain responsible for developing the government policy on fight against climate change, prevention of pollution, protecting the natural heritage, biodiversity, forests, sea, water, and energy for a more ecological and productive social model. Likewise, it is responsible for the elaboration and development of the government policy against the country's demographic challenges (population ageing, territorial depopulation, floating population effects, etc.). Within MITERD, the Spanish Inventory System (SEI) is the specific Spanish facility in charge of the National Inventory of emissions and removals of greenhouse gases and atmospheric pollutants, as well as Projections of emissions and removals into the atmosphere, which allow assessing compliance with the commitments assumed by Spain within the framework of international and European regulations on emissions into the atmosphere. Likewise, they are the basis for the development of emission mitigation policies and measures and for the evaluation of their effectiveness in achieving the proposed objectives in compliance with its information obligations.

Following relevant guidelines and methodologies, the Spanish Inventory System (SEI) annually prepares the Inventory of Greenhouse Gas Emissions by anthropogenic sources and their absorption by sinks, as provided in the Framework Convention on Climate Change (UNFCCC) and its Kyoto Protocol and in the Regulation (EU) 525/2013 for the monitoring and notification of greenhouse gas emissions and other information relevant to climate change. The methodologies used to estimate emissions in each sector of activity, including LULUCF, are described in Sector Files, which are currently being prepared and published in the web of MITECO (www.miteco.gob.es). The Sector Sheets are technical documents that provide a generic description of the methodologies used to estimate emissions in order to give the process as transparent as possible. Each File is published individually, indicating its date of elaboration. The eventual updates of the Sheets are independent of the calendar followed by the Official Reports, so their content and, especially, the data in their Annexes, may not coincide with those of the latest edition of the Inventory, which prevails over the Sector Sheets. As additional information, the old methodologies used for the 1990-2012 Series of the National Inventory of Atmospheric Emissions (2014 Edition: Volume 2) are still available for consultation. However, these methodologies are currently obsolete, having been modified in successive editions of the National Inventory. The Inventory of Emissions to the Atmosphere is included in the National Statistical Plan 2021-2024, elaborated following a standardized methodological report (IMS) also available in the web of MITERD.

The QA/QC plan in Spain has been designed following the guidance provided by the IPCC 2006 Guidance and the EMEP/EEA 2019 Guidance, using additionally as reference the European Commission's SWD (2013) 308. Although most of the data providers for the SEI have their own QA/QC systems, the Spanish Inventory System coordinates and complements QA/QC activities to meet its own quality objectives. The QA/QC system of the National Inventory tries to harmonise the availability of time and resources and uses the PDCA (plan-do-check-act) approach. The General Directorate for Environmental Quality and Assessment and the Natural Environment (DGCEA), as the competent authority of the Spanish Inventory System (SEI), is the body responsible for the QA/QC system of the National Inventory, acting as the QA/QC administrator, and is supported by specific technical assistance to carry out the necessary tasks.

Throughout the National Inventory cycle, different quality control activities and procedures are carried out and documented, these activities being ordered according to the stage of the National Inventory cycle in which they take place. There are four main groups of controls, which are listed below:

- Controls on the implementation of improvements
- Controls over the collection of information
- Controls over the processing and loading of data into the database
- Controls on emission results

In addition, during the preparation of reports, each technician responsible for drafting a section or chapter uses and fills in a checklist. In addition, a second person, as an additional quality control, reviews each chapter of the report.

The quality control and documentation tools of the National Inventory include the improvement plan, the database of information requests, the data import tool, the quality control tool, the checklist for the preparation of reports, the Spanish Emission Inventory Quality Management Tool (HGCIEE) and the recalculation analysis tool.

On the other hand, the quality assurance (QA) system includes a series of activities performed by third parties, which are not directly involved in the development process of the National Inventory. These activities are aimed at verifying compliance with the reporting requirements and assessing the effectiveness of the QC system. Reviews of the national GHG and air pollutant inventories are conducted annually. The main outcome of these reviews is a list of issues and recommendations that are incorporated into the National Inventory improvement plan. In addition, since October 2017, an independent consulting firm (IDOM Consulting, Engineering, Architecture SAU) has been carrying out a quality assurance audit of the National Inventory.

The National Emissions Inventory includes the result of estimating the emissions and removals associated with the changes in C stocks from deposits of the six land use categories as well as the additional deposit, wood products. In addition to emissions / removals related to variations in C deposits, in the LULUCF sector can be estimated:

1. direct N_2O emissions from nitrogen (N) inputs in managed soils;
2. emissions and removals from drainage and rewetting and other organic soil management practices and minerals;
3. direct N_2O emissions from nitrogen (N) mineralization related to loss of matter organic in mineral soils due to changes in land use;
4. indirect N_2O emissions from N leaching and runoff related to the loss of organic matter in mineral soils due to change in land use; and
5. greenhouse gas emissions from fires and controlled burning, as well as atmospheric pollutants.

The National Emissions Inventory includes estimates for the last three, as it considers that the first two does not take place in Spain. Annex I of this methodological sheet includes two tables that summarize the coverage of the estimation of emissions and removals associated with the LULUCF sector.

In addition, National System Inventory of LULUCF sector provides information on the following topics:

1. Introductory sheet to the Land Uses, Land Use Changes and Forestry (LULUCF) sector, as explained above
2. Contributions of N in managed soils
3. Drainage and rewetting and other organic and mineral soil management practices
4. N mineralization due to loss of soil organic matter due to changes in land use in mineral soils
5. Leaching and runoff of mineralized N due to loss of soil organic matter due to changes in land use in mineral soils
6. Change in carbon stocks from living biomass on remaining forest land
7. Change in carbon stocks of deadwood in transition lands
8. Change in soil organic carbon stocks in mineral soils in transition lands
9. Change in carbon stocks in wood products

Currently, Spain uses a combination of **databases for the reporting of the LULUCF sector,** that includes:

1. The forest map of Spain (MFE), with the available cartographic products of the years 2009, 2012 and 2015,
2. The crop and land use maps of Spain (MCA);

3. The Geographic Information System for Agricultural Parcels (SIGPAC) 2009, 2012 and 2015.
4. Corine Land Cover (CLC) 1990, 2000, 2006, 2012 and 2018.

The procedure used in the National Inventory Report to estimate the areas of land uses and land use changes in the period 1990-2018 is based on the exploitation of different cartographic and statistical databases available at country extent (*Figure 1.8*) on which an adjustment was applied statistically for land afforestation.

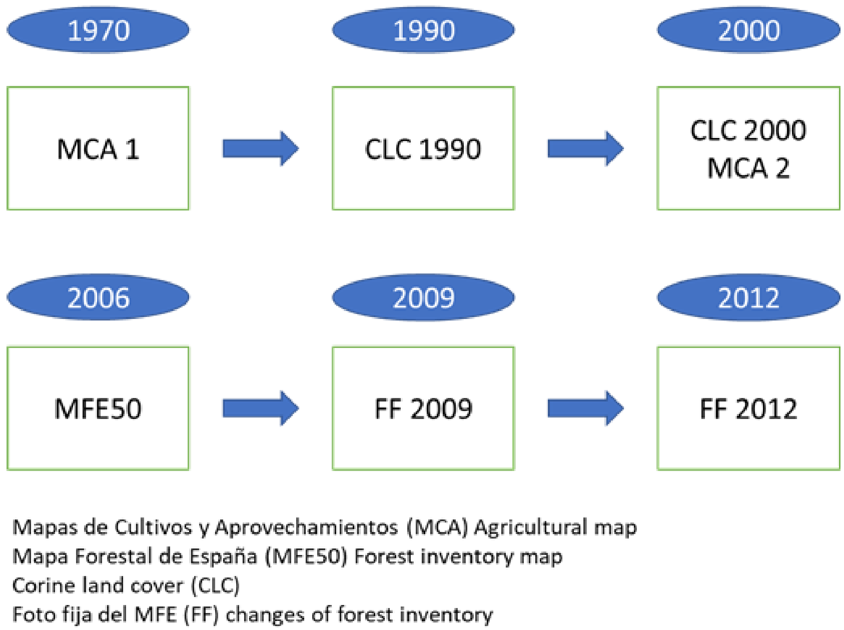


Figure 1.8: Data sources used by Spain to calculate the area information by cartographic approach in its NIR 2019 for LULUCF sector.

The cartographic exploitation includes the datasets for the period 1989-2005. One step beyond, the namely "foto fija" 2009 and 2012 were included to track the transition from forest lands to croplands, wetlands and settlements. This cartographic information database was processed with the tools of a geographic information system. In addition, the 2018 edition of the National Inventory included the provisional estimation of land use areas and changes in land use for the period 1970-1989, based on available statistical information.

The forest map of Spain (MFE), serves as base cartography for the National Forest Inventory, being the basic forestry cartography at state level, which shows the distribution of the Spanish forest ecosystems. This map uses a working methodology based on photo-interpretation, with field verification. The MFE provides detailed and homogeneous vector information for the whole Spanish territory on the structural type or main use of each plot, the degree of cover and the main tree species mapped, among others. The most current MFE is the MFE at a scale of 1:25.000 (MFE25), although it is not available for all the provinces of the Spanish territory, so for those where it is not available the MFE at a scale of 1:50.000 (MFE50) is used.

The Map of Crops and Use of Spain (MCA)n starts in the 1970s, with the collection of Crop and Use Maps at a scale of 1:50,000, which was and still is the only existing cartography. At this scale and for the entire national territory reflects the distribution of land use in Spain. This cartography has been used since then and continues to be used for numerous agricultural and environmental studies and projects. This MCA offers cartographic and alphanumeric information in three sections: Map of land uses and overloads, Report of uses and overloads, (both by complete sheet and by municipality) and Auxiliary maps of slopes and altitudes (according to sheet 1:50,000). All this information can be consulted for the entire national territory.

The Geographic Information System for Agricultural Parcels (SIGPAC), developed for the years 2009, 2012 and 2015, allows the geographical identification of parcels declared by farmers in any area-related aid scheme. The system consists of a mosaic of digital orthophotos covering the entire national territory, on which the land registry maps are initially superimposed, so that for each specific reference, the system automatically provides the image of the referenced parcel on screen. Among the most important objectives of the SIGPAC Project are to facilitate the submission of applications by farmers, to facilitate administrative controls and to facilitate on-the-spot controls, speeding up the location of plots and allowing "quick visits" in both classic and remote sensing controls.

Finally, Spain uses the Corine Land Cover (CLC). The CORINE Land Cover (CLC) inventory was initiated in 1985 (reference year 1990), with new updates in 2000, 2006, 2012, and 2018. It consists of an inventory of land cover in 44 classes. The time series are complemented by change layers, which highlight changes in land cover. The majority of countries produce CLC by visual interpretation of high-resolution satellite imagery. However, in a few countries semi-automatic additional solutions are applied, using national in-situ data, satellite image processing, GIS integration and generalisation. It has a wide range of applications, on which various Community policies are based in different domains, such as agriculture, transport, spatial planning, or the domains of environment.

This heterogeneous pool of maps and databases is currently being managed towards creating a cartographic system fulfilling an IPCC Approach 3 (geospatially explicit). At this regard, the Spanish LULUCF inventory has recently made an effort to develop a geographically explicit land use change monitoring system according to IPCC Approach 3 (geospatially explicit) from 1970 to the present. This system has been based on the collection and homogenization of all the existing historical cartography of land use and land cover (forestry, crops, CORINE, SIGPAC, etc) from 1970 to 2015. The results of this new project are going to be used in the calculation of the emissions and removals of the LULUCF sector in the next 2020 edition of the National Emissions Inventory. One step beyond, the National Inventory System is currently evaluating/exploring different alternatives for including the temporal dimension into the actual geographically explicit land use-land cover change monitoring system, for the foreseen 2015-2030 monitoring period. Among the possible options, this project has fostered assessing the capabilities, tools, and datasets from the Copernicus program.

The combination of Regional or Member State-specific data needs for monitoring emissions and removals from land use, land use change and forestry, together with the existing and forthcoming Copernicus data and services will be an opportunity to evaluate the data coherence and compatibility that may exist at a national level between historical cartography and Copernicus products. Added to the evaluation of such degree of compatibility, another result of the project has been the identification of the main obstacles and gaps, and so will be considered the technical solutions necessary for a robust and coherent articulation between the two systems.

In addition, the Spanish inventory is seeking a Tier 2 (T2) in the categories cropland (CL), grassland (GL) and wetland (WL) in the Mediterranean domain, providing also consistency with SIGPAC database. The Spanish LULUCF Inventory has a recognized experience in the application of Tier 2 methodological levels with "country specific" data for certain use transitions or even remaining uses (i.e. Cropland remaining cropland for the case of transitions involving a woody crop and particularly olive groves and vineyard). However, for the uses of CL, GL and WL that represent an area of 323,000 km2 (approximately 64% of the national territory), Tier 1 methodologies are often applied for the estimation of emissions and absorptions (E/A) linked to changes in live biomass (LB), and organic carbon in the soil (SOC) in the remaining uses, assuming a neutral balance. In this sense, Copernicus tools are considered to be able provide very useful information regarding management practices of these soils, or the temporal evolution of their natural state that could allow better inferred carbon gains or losses by applying Tier 2 methodologies.

The project development will help analysing different technical options to improve the E/A estimations using T2 methodologies in uses such as CL, GL, and WL. In addition, such analysis at the Spanish national level would imply the evaluation of different climatic areas (Mediterranean, humid temperate or insular), with their intrinsic characteristics. The technological solutions that would be proposed as a result of the project could be extrapolated to other countries in the Mediterranean area.

The methodology used in the LULUCF sector follows the guidelines of the 2006 IPCC Guide, and, in part, of the 2013 Wetlands Supplement, of the 2013 KP Supplementary Guide and of the 2003 IPCC Good Practice Guide, using, in the emission / removal estimation algorithms, national parameters whenever possible, while in cases where such information is not available, those proposed in the 2006 IPCC Guide. Spain is currently working on the production of a consistent time series on LULUCF maps to comply with the EU Regulation 2018/841. This new generation of maps will correspond to Approach 3 of IPCC 2006 guidance, as explained. According to the information provided by the Ministry in the NIR 2020, these new maps will be implemented in the next NIR.

After a detailed process of engagement with final users (SEI and all related MITERD departments) and a total of 13 thematic meetings, **existing gaps** were identified. They are the following:

Forest lands: data related to afforestation / reforestation (occurrence, species identification and monitoring success), identification of natural disturbances (disease infestations, insect pests, extreme weather events and geological disturbances) and identification of timber and fuelwood harvests (thinning, pruning, cutting, wood removed and quantification of the extracted biomass).

Croplands: data related to the soil organic carbon (SOC) in agricultural crop (estimation and monitoring), discrimination of rainfed and irrigated crops, Identification of fallow lands (Time lying fallow), identification of herbaceous CL management practices (sowing, tilling, irrigation, fertilisation, pruning and area affected by this practise), identification of woody CL management practices (tilling and irrigation) and identification of crop and pruning waste management system (stay in the field without burning, field burning and removal).

Grasslands: data related to vegetation maps (improved mapping for canopy cover<20%) (woody tree, woody shrub and herbaceous), quantification of carbon stock (living aerial biomass, underground living biomass, deadwood, detritus and soil organic carbon and organic carbon content), discrimination of pastures by management type (tilling, irrigation, fertilisation and species improvement) and discrimination of grassland by use (mowing, grazing and intensity of use).

Wetlands: data related to improvements in the identification of wetland uses (fish farms, salt pans, marshes, etc.), discrimination between wetland types (lakes, salt marshes, marshes, reservoirs, ponds, basins, fish farms, rivers, canals and ditches) and monitoring of processes related to water and wet systems.

Stakeholders for national reporting data are, as explained above, the Ministry for the Ecological Transition and the Demographic Challenge (MITERD), specifically the Spanish Inventory System (SEI) department.

As stated before, the project development is expected to help analysing different technical options to improve the E/A estimations using T2 methodologies in uses such as CL, GL, and WL. In addition, such analysis at the Spanish national level would imply the evaluation of different climatic areas (Mediterranean, humid temperate or insular), with their intrinsic characteristics. The technological solutions that would be proposed as a result of the project could be extrapolated to other countries in the Mediterranean area. Also, the aim of this project is to obtain a series of CLMS solutions and remote sensing solutions in general that can be used to assist MITERD in the LULUCF report. In this way, by combining these solutions with the data already available to MITERD, a suitable methodology for future National Emissions Inventory reports can be obtained.

2. European Copernicus data for national LULUCF reporting

Pan-European High-Resolution Layers (HRL) provide information on specific land cover characteristics and are complementary to land cover / land use mapping such as in the CORINE land cover (CLC) datasets (*Table 2.1*). The HRLs are produced from satellite imagery through a combination of automatic processing and interactive rule-based classification. Since the 2015 reference year, the production is increasingly based on time series of satellite images from several different sensors, including the combination of optical and radar data. The main sources are now (since the 2018 reference year) the Sentinel Satellites (particularly Sentinel-2 and Sentinel-1). In addition to high resolution (HR) data, since 2015, we also use very high resolution (VHR) imagery for some of the products. Since 2018, the products have increased in resolution to 10 meters, thus following the source resolution of the Sentinel-2 imagery.

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. All these five products are continuing existing products, some with longer time series existing (Imperviousness and tree-cover/forest), and three products that have only one previous reference year (2015) (grassland, the water & wetness products and Small Woody Features). All products are mapping the features under consideration for the whole of the EEA-39 area. They are produced in a combined centralized and decentralized approach, involving service industry through market mechanisms and participating countries through grant agreements.

The HRLs can then be used, for example, as attributes for different kind of more aggregated spatial units, such as NUTS3, CLC polygons, regular grids or designated areas.

Table 2.1: Pan-European High-Resolution Layers (HRL) on specific land cover characteristics.

	2012 production	**2015 production**	**2018 production**
Imperviousness	Imperviousness and imperviousness change for reference years 2006, 2009, 2012 and change products in 20m resolution	Imperviousness Degree and imperviousness change. Full re-processing for reference years 2006, 2009, 2012 and change products, and 2015 status products in 20m resolution	Imperviousness Degree (IMD) for reference year 2018 in 10 meter resolution , imperviousness change (IMC) and imperviousness change classified (IMCC) at 20 m resolution . Addition of Impervious Built-up (IBU) at 20 m resolution and its corresponding 100 meter aggregate Share of Built-up (SBU).
Forest	Tree cover density, Dominant Leaf Type and Forest Type products for reference year 2012 in 20m resolution	Tree cover density, dominant leaf type and forest type products + new change product (Tree Cover Change Mask) for reference year 2015 in 20m resolution	Tree cover density, dominant leaf type and forest type products for reference year 2018 in 10 m resolution; Tree cover change mask (TCCM) and Dominant Leaf Type Change (DLTC) at 20 m resolution.
Grassland	n.a.	New grassland baseline product, including all grasslands for reference year 2015 in 20m resolution	Grassland status product for reference year 2018 in 10m resolution. Addition of new 2015-2018 Grassland change in 20m resolution.
Wetness and Water	n.a.	New combined baseline product based on 7-year time series (2009-2015) analysis mapping temporary and permanent wet and temporary and permanent water status	Water & Wetness status product based on 7-year time series (2012-2018) analysis, mapping temporary and permanent wet, and temporary and permanent water status

		for reference year 2015 in 20m resolution	for the reference year 2018 in 10 meter resolution. The update uses a "rolling-archive" approach with overlapping 7-year time periods.
Small Woody Features	n.a.	New product based on VHR data, mapping small patchy (200 m² ≤ area ≤ 5000 m²) and linear woody features as vector product, but also available in 5m and 100m raster version.	Later addition

S2GLC land cover classification approach has been developed in the frame of ESA project - Sentinel2 Global Land Cover (S2GLC). The developed classification is a pixel-oriented, supervised approach in which all processing steps are performed automatically without a manual intervention like an interpretation of satellite or reference/training data. Training points are selected using existing land cover databases which can be prepared even with very low resolution compared with classified satellite image. This solution making it a very effective tool for LC classification of the whole countries or continents. Initial classifications are performed on separate tiles of Sentinel-2 images, collected throughout the year, using the Random Forest (RF) classification method. These individual results are then combined using a developed aggregation approach. The final classification result of each pixel is determined by reviewing the prediction scores of all individual classifications. The proposed aggregation method performs better than processing sets of data using functions dedicated to time series, particularly in cases where available images contain clouds. The last step in the classification is post-processing, to improve the results of the land cover map. All classification steps, processing of satellite and auxiliary data are performed using software developed by CBK PAN, the final LC resolution is 10x10 m. (Lewinski et. al. 2017, Malinowski et. al. 2020). S2GLC classification approach was used for LC classification of Europe (*Figure 2.1*). Over fifteen thousand Sentinel-2 images, representing 815 Sentinel-2 tiles collecting in the year 2017 have been processed. Validation of the final map was performed based on a large set of 52 000 randomly distributed samples representing 55 Sentinel-2 tiles spread across Europe. The overall accuracy (OA) of the complete map with 13 LC classes was estimated to be over 86%. Due to lower accuracy achieved for recognition of some of the classes (e.g. grasslands and moors), an additional merging of selected vegetation classes was proposed. It resulted in a reduction of the number of LC classes to 10 but assured an increase of OA up to 89%. The accuracy assessment on a country level revealed the very good quality of the LC map with most countries exceeding 80% of OA. The obtained results should be assessed as very good as for the LC classification approach performed with so high degree of automation. The S2GLC approach is improved, in 2020 on the order of Polish Space Agency classification, Poland LC 2020 was performed.

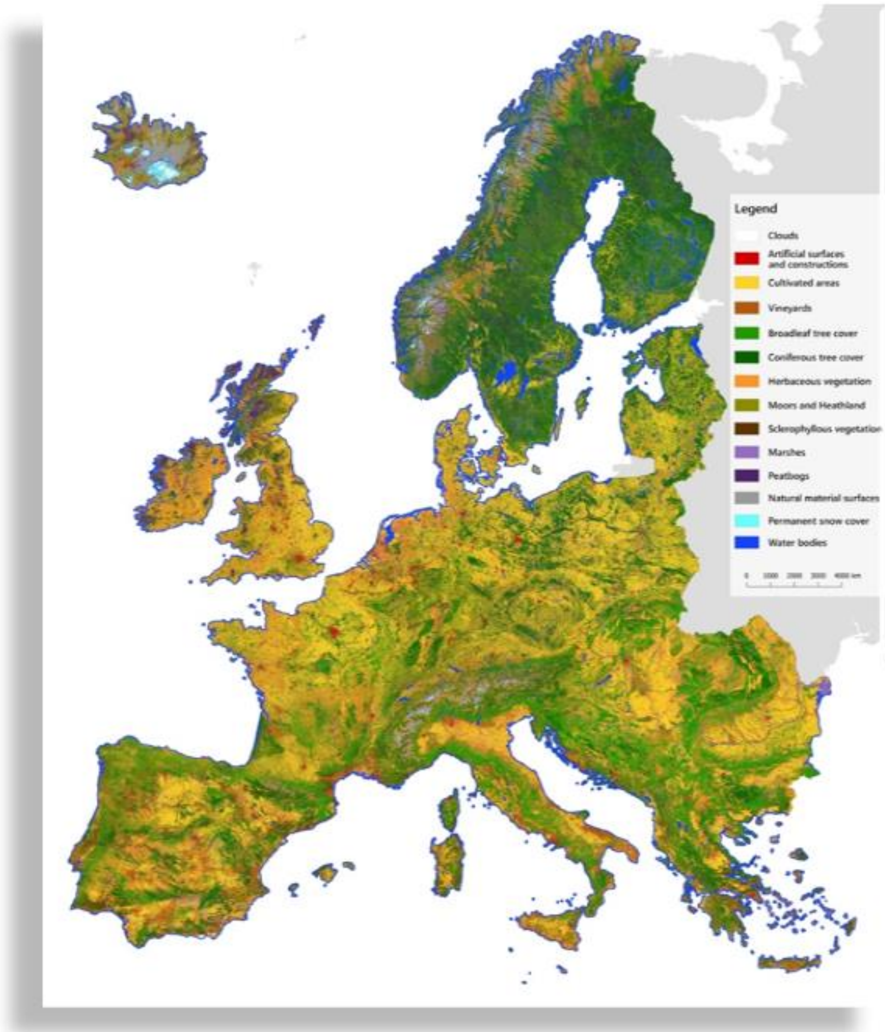


Figure 2.1: LC Europe 2017 – land cover classification performed using S2GLC approach.

European Corine Land Cover data has been produced since 1990 and from 2000 onwards every six years. In majority of countries the CLC is produced by visual interpretation of high-resolution satellite imagery. In some countries also semi-automatic methods are applied such as satellite image processing, GIS integration and generalization. The data consists of 44 land cover classes with MMU of 25 hectares for areal and minimum width of 100 m for linear phenomena.

Maynooth University used Google Earth Engine to classify land cover from 2000 – 2020. The COPERNICUS products that were used in this process included Sentinel 2 data for images to run classification techniques. Land Use/Cover Area frame Survey (LUCAS) was used to validate the image classification. As the CORINE land Cover inventory is used in the Irish national inventory to estimate wetlands, settlements Grasslands and forestry, an accuracy assessment is an important step in validating national inventory arrangements for the LULUCF sector. As validation is an important component in achieving transparency, accuracy, completeness, comparability and consistency (TACCC), the application of LUCAS to provide metrics on accuracy may be an important improvement over existing validation measures.

Table 2.2: Copernicus Products used with applications

Product	**Application**
Sentinel 2	Image Classification, development of timeseries
CORINE	To assess how appropriate CORINE is for inclusion in LULUCF inventory
LUCAS	Validation of Image Classification
HRL	To assess how appropriate HRL is for inclusion in LULUCF inventory

Charles University used Google Earth Engine to classify land cover from 2017 – 2020. The COPERNICUS products that were used in this process included Sentinel 2 data for images to run classification techniques of Random forest techniques. An accuracy assessment is an important step and used in-situ data and aerial photographs.

SRTI-BAS tested two approaches for land cover/land use mapping which differed by the input data used. In the first approach satellite imagery from Landsat 7, Landsat 8, Sentinel-1, and Sentinel-2 were used to run random forest classification from 2012 to 2018. Training data and test data (for accuracy assessment) were collected by means of visual interpretation of high-resolution imagery from Google Earth. The "Settlement" class was mapped using the CORINE Land Cover 2018 dataset. In the second approach Copernicus High Resolution Layers (HRL), namely Imperviousness Density (IMD), Tree cover density (TCD), Dominant leaf type (DLT), Grassland (GRA), and Water and wetness (WAW) were combined to produce land cover maps for 2015 and 2018. However, the HRLs data were not enough and had to be complemented with other data. For example, CORINE Land Cover 2018 dataset was used to better map the settlements, while Sentinel-2 derived land cover information extracted either by image classification or vegetation index thresholding was used to map areas not covered by the HRLs.

IHCantabria and ETC-ULS-UMA (a subcontractor) are currently doing a detailed assessment of the Copernicus Land Monitoring Service (CLMS) and its potential contribution to spatially-explicit LULUCF reporting in the country, by evaluating independently the CLMS portfolio and CLC+ advances for LULUCF regulation and selecting representative examples of this support.

The recently adopted EU LULUCF Regulation (2018/841) will require MS to use geographically explicit land-use conversion data for the accounting of emissions and removals from the LULUCF sector for the years 2021 to 2030. In addition, the EU LULUCF Regulation explicitly outlines the role of the EEA in assisting the European Commission with implementing this legislation. Copernicus Land Monitoring Services (CLMS) like Corine Land Cover (CLC) and its successor, CLC+, which is currently under development are spatially explicit products and may therefore serve the MS for future LULUCF reporting and/or the quality assurance of the reported land use and land use change matrices. Several Member States, as Spain, already use CLC for LULUCF GHG reporting because it is spatially explicit and it allows identifying land cover and land use changes for a large part of the required time series for LULUCF reporting, which needs land use change data since 1971. This is done together with other national and/or international sources of geographical data for compiling their LULUCF inventories.

To help evaluate the potential of CLMS products to support future LULUCF reporting, a study comparing the specifications of these products against the explicit land use category definitions and stratifications used by Spain is being carried out. This study focussed on:

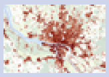
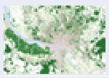
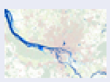
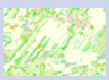
1. How Spain defines the IPCC land use categories, as well as how these categories are further stratified. Stratification is understood as the practise of splitting the main category between certain geographical strata and applying respective stratum-specific C stock values and/or emissions factors to calculate GHG removals/emissions;
2. What is the current methodology used by Spain to monitor the LULUCF sector under the national inventory report (NIR);
3. What extent the products CLMS currently offered can be used for LULUCF reporting and which factors may be preventing further use of these products. Ultimately this will inform the development of a CLC+ LULUCF instance to support the LULUCF GHG emissions reporting as of 2023.

As part of the assessment of state-of-art and capitalization of EU projects and actions related to the implementation of the EU LULUCF regulation, the work developed by ETC on Urban, Land and Soil Systems (ETC/ULS, see https://www.eionet.europa.eu/etcs/etc-uls) in support to EEA, was reviewed. In 2019, ETC/ULS developed a task (ETC/ULS task 1.4.4.1 – AP2019) that surveyed categories and strata as defined by Member States and analysed similarities, deviations, challenges in relation to the potential CLMS uses/solutions in relation to the demands of using geographical explicit land use data.

The ETC/ULS' work in relation to LULUCF is not freely accessible although was presented to LULUCF and Copernicus community in 2019. The ETC/ULS team offered the 2019 final report to capitalize EU efforts. We provide here some examples.

CLMS products and services are made available, free and open through the Copernicus land portal https://land.copernicus.eu/ both as view service, web map services and for download. In turn, HRLs (*Table 2.3*) provide information on specific land cover characteristics and are complementary to land cover / land use mapping such as in the CORINE land cover (CLC) datasets. They are based on HR and VHR imagery and have 20m resolution (10m from 2018 on, following Sentinel-2 imagery), giving coverage for EEA 38 + UK with a set of sub-products, including the systematic monitoring of biophysical parameters that consists of a set of bio-geophysical products on the status and evolution of the land surface, such as Vegetation Phenology and Productivity from 2017 at the level of Vegetation Index (LAI, FAPAR, NDVI, PPI) (Near-real time), Seasonal Trajectories (every 10 days) and Vegetation Phenology Parameters (yearly) with a spatial resolution of 10 meters.

Table 2.3: HRL products.

HRL	Subproduct	Time series	Change layer	Resolutions
Imperviousness 	Density (% of sealed area) Built-up	2006, 2009, 2012, 2015, 2018	2006-2009, 2006-2012, 2009-2012, 2012-2015, 2015-2018	20m & 100m 10m & 100m (2018)
Forest 	Tree cover density (0-100%) Dominant leaf type (broadleaved vs. coniferous) Forest Type product (=> FAO definition)	2012, 2015, 2018	Tree cover Change 2012-2015, 2015-2018 DLT change 2015-2018	20m (2012, 2015) 10m (2018)
Grassland 	Status (binary) Probability (0-100%) Ploughing indicator (nº of years)	2015, 2018	2015-2018	20m (2015) 10m (2018)
Water and Wetness 	WAW classes (permanent, temporary) Probability (0-100%)	2015, 2018		20m (2015) 10m (2018)
Small Woody Features 	Linear, Patchy features Density layer (0-100%)	2015		5m, 100m

The CORINE Land Cover map is for land monitoring in Europe. This dataset has a number advantages:

1. Consistent timeline since 1990
2. Widely recognized
3. Good overview of Europe
4. Relatively easy production implementation on country level (44 classes)
5. Defined quality of these datasets

But also some issues and caveats:

1. Nowadays too coarse (MMU with 25 ha)
2. Different MMU for change layer (5 ha)
3. Mixture of land cover and land use
4. One class for each polygon: no additional thematic attribution
5. 6-year production interval

This has led to the definition of CLC+ as the suite of products under development of the CLC 2nd Generation approach with its CLC+ Backbone (production started Q2 2020), CLC+ Core (production started Q4 2020), CLC+ Instances and CLC+ Legacy. As a conclusion, evolving policy requirements need higher resolution, more frequent and more flexible land use & land cover information. Next generation Corine Land Cover (CLC+) is one set of tools to help get there, moving from classification to characterisation/description, moving from lower spatial and temporal resolution to higher spatial and temporal resolution with a grid based data-cube allows flexible "on demand" products (instances).

In Spain, Forest land (FL) is one of the categories that has been evaluated more by comparing CLMS data and products with currently available Spain forestry map MFE. When comparing with HRL we find a number of advantages;

1. Time series available but from 2012 on, not possibility to time-back easily
2. No Minimum Mapping Unit (MMU); pixel-based at 20m spatial resolution (2012 – 2015) and at 10m (2018)
3. TCD provides continuous information (0 – 100%) of crown cover. This allows to define the threshold for setting the national forest definition.

But also, limitations:

1. Min height has not a clear information
2. Time lapse of 3 year
3. The change products, TCDC, cannot be used directly for mapping the use. The product provides information of increase / decrease of density of tree cover.
4. The HRLs are sectorial products (HRL forest, HRL grassland, HRL imperviousness, etc.). Then, it would need a composition of multiple datasets to identify the land covers involved in each change (from / to).
5. Some products are pure cover products; the permanent wood crops are included in these products. This is an important issue regarding the clear-cut harvests parcels that remain as Forest Land. However, in order to separate real “forest” areas from non-forest areas, an additional raster product is provided by CLMS named Forest Additional Support Layer (20m). The improved product id the FTY that exclude tree under agri and urban.

When comparing to CLC, we also find some possibilities:

1. Time series available, consistent in time from (1990) 2000 to 2018, every 6 years
2. It is partially a Land use /land cover product.

But exist many caveats (we can see in *Figure 2.2*):

1. MMU in CLC is 25 ha/ 5ha for change
2. Minimum height is 5m. This could be an issue when the country specification is lower
3. Minimum crown cover is 30%. This could be a problem for countries with lower thresholds. However, in most of the forest related actions (KIP-INCA, MAES, LEAC), the class 3.2.4 (transitional forest) is included as a forest area.
4. Time lapse is 6 years
5. Time consistencies 2018 onward not available, although CLC legacy can be represent an opportunity.

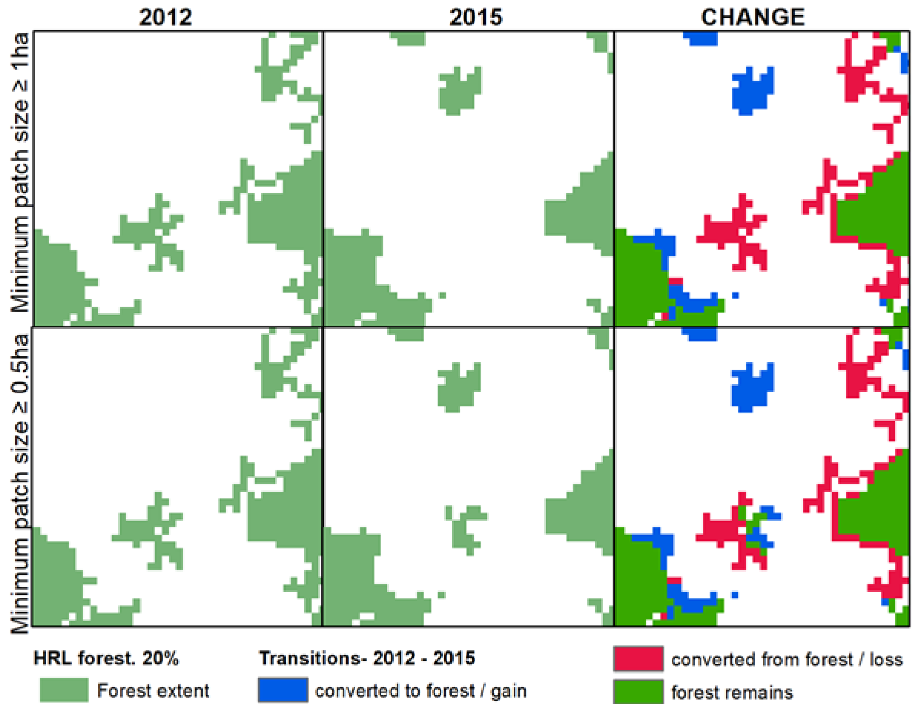


Figure 2.2: Example to show the effect of forest land categories threshold on the detection of changes.

This task will provide a complete review of all capabilities of CLMS data and services for LULUCF reporting in Spain. It is still ongoing and as planned to be finished and reported at the endof May 2021.

Finland compared the areas of IPCC land use classes as reported in the national GHGI to areas produced from various CLMS datasets. These were

1. Finnish national high resolution Corine Land Cover classification (20m raster) (**FI CLC18**). National high resolution Corine Land Cover (Finland) has been produced every six years since year 2000 (and thus is available for 2000, 2006, 2012 and 2018). It is a raster data with resolution of 20x20m (25x25m for the first 2 production years). Change layers in raster format have been produced for 00-06, 06-12 and 12-18 with MMU of 0,5ha. Unlike the EU Corine land cover vector data, the Finnish national dataset is classified to 4th level with 44-49 classes, depending on the production year. The HR CLC is semi automatically interpreted from satellite imagery (Landsat-7, Spot, IRS LISS, Sentinel-2 depending on time period) and then further enriched with national land use data. Relevant national datasets used in the production are:
    a. Agricultural Land Parcel Identification System
    b. Topographic database
    c. Building and Dwelling Register
    d. Digiroad
    e. National forest inventory
    f. Database on water bodies and river beds
    g. Registers on dump and mineral extraction sites and peat production areas
    h. Some land use classes have been digitized by hand, such as golf courses, dump sites, harbours, open cast mines and large construction sites.
2. Global Land Cover produced from Sentinel-2 images by Polish academy of Sciences (**S2GLC**)

3. Global Land Cover produced from Sentinel-2 images by Polish academy of sciences + CLMS High resolution Layers (**S2GLC & HRLs**)
4. CLMS High Resolution Layers + European Corine Land Cover with 25 ha (**HRLs & EU CLC**)

Table 2.4: Classification methods of Copernicus based land use data into IPCC-classes

IPCC-class	**Fi CLC18**	**S2GLC**	**S2GLC & HRLs**	**HRLs & EU CLC**
1. Forest land	3111, 3112, 3121, 3122, 3123, 3131, 3132, 3133, 3241, 3242, 3243, 3244	82 Broadleaf tree cover, 83 Coniferous tree cover	82 Broadleaf tree cover, Coniferous tree cover OR HRL TCD > 30%	HRL Forest TCD >= 30% & <= 100% (3)
2. Cropland	2111, 2221, 2311	73 Cultivated areas, 75 Vineyards	73 Cultivated areas OR 75 Vineyards	EU CLC == 211
3. Grassland	2312, 2431, 2441	102 Herbaceous vegetation	102 Herbaceous vegetation	HRL Grassland
4. Wetlands	4111, 4121, 4211	105 Marshes, 106 Peatbogs	105 Marshes OR 106 Peatbogs	WAW 2-4 ORI EU CLC 411, 412, 421
42. wetlands, peat extraction	4122	-	-	-
43. Wetlands, inland water	4112, 4212, 5111, 5121, 5231	162 Water bodies	162 Water bodies OR HRL Permanent Water	HRL Water (WAW == 1)
5. Settlements	1111, 1121, 1211, 1212, 1221, 1231, 1241, 1311, 1312,	62 Artificial surfaces and constructions	62 Artificial surfaces and	HRL Imperviousness >= 30%

	1321, 1331, 1421, 1422, 1423, 1424, 3246		constructions OR HRL Imp > 30%	
6. Other land	3211, 3221, 3311, 3321, 3331	103 Moors and Heathland, 104 Sclerophyllous vegetation, 121 Natural material surfaces, 123 Permanent snow covered surfaces	103 Moors and Heathland OR 104 Sclerophyllous vegetation OR 121 Natural material surfaces OR 123 Permanent snow covered surfaces	The rest

Foreseen potential of Copernicus products in LULUCF reporting in Finland were tested by using the National High resolution Corine Land Cover datasets from 2000, 2006, 2012 and 2018. Emissions and removals in GHGI reporting categories were calculated using the areas based on this data. The CORINE data has information on mineral and organic soils. In a case where the information on soil type was missing, the Finnish soil database was used. SYKE provided this data for Luke for the analysis.

Land cover and use classifications that SYKE provided for FMI for the atmospheric top-down approaches were based on

1. European Corine land cover 2018, 2012 & 2006
2. European Corine land cover 2018, 2012 & 2006 and European peatland map (Tanneberger et.al., 2017)
3. S2GLC & CLMS HRL Imperviousness, Forest TCD, Water
4. Finnish national Corine Land Cover 2018, 2012, 2006

LC classes of these classifications were

1. Settlement
2. Agricultural lands
3. Forest land, mineral soil
4. Forest land, peat
5. Forest land, afforested mineral soil
6. Forest land, afforested peat
7. Transitional woodland (tree crown cover 10-30%), mineral soil
8. Transitional woodland, peat
9. Transitional woodland, deforested
10. Transitional woodland, deforested
11. Open mineral soil (tree crown cover < 10%)
12. Wetlands, Marsh

13. Wetlands, Open bogs
14. Wetlands, peat production
15. Water (sea, lake, river)

3. Proposed solutions and methodologies for improving the current existing LULUCF methodology and regulations

3.1. User engagement

Maynooth University (MU) have had 3 meetings with the inventory agency where a project was formulated based on identified needs and requirements with the inventory agency.

The inventory agency and MU developed a strategy where MU can provide technical assistance to the inventory agency in research and development in integrating beneficial years in the land use inventory, gap filling procedures and the provision of indicative maps for 1990.

Several other meetings are planned to update the inventory agency and facilitate data exchange. In addition, a workshop is planned for July 2021 and may include several stakeholders ranging from Teagasc (Irish Agricultural Research Institute), Bord Na Mona (The national Peat Bord), Department of Agriculture and others to be identified later.

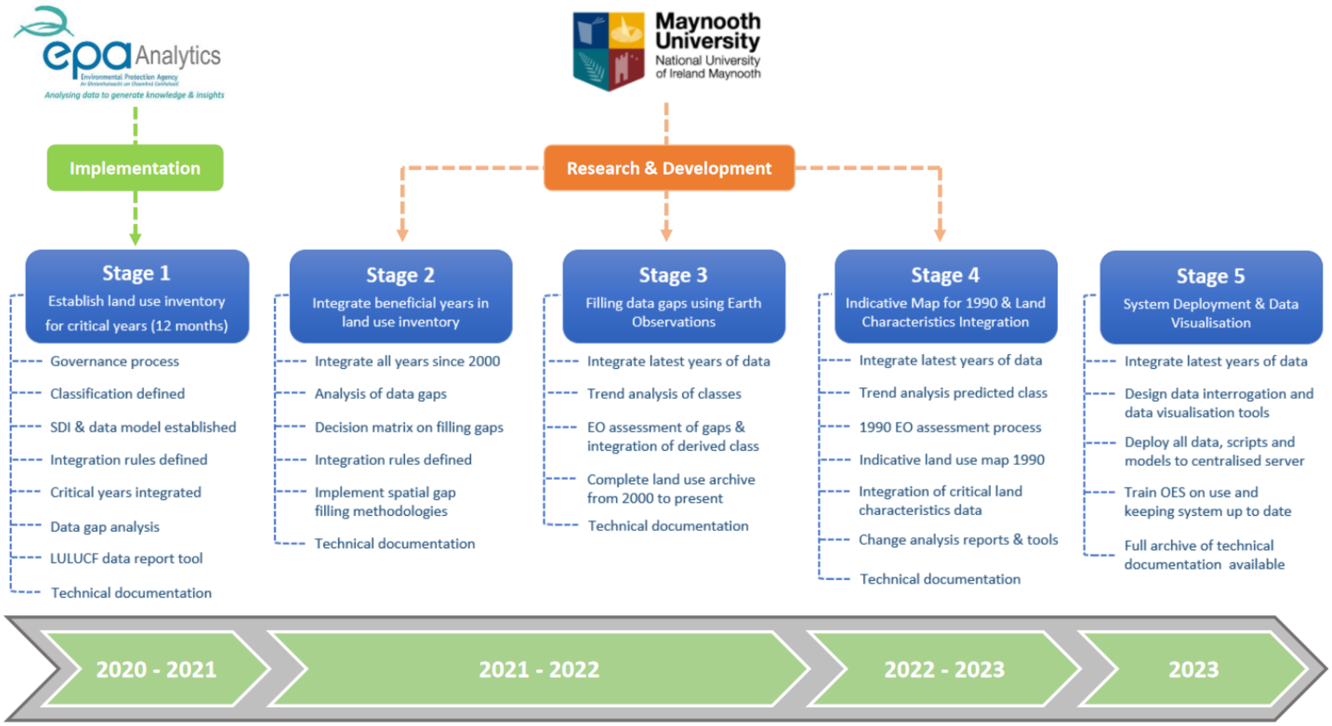


Figure 3.1: Inventory Agency research and development plan and Maynooth University's potential role in supporting inventory.

Bilateral meetings were organised with Statistics Finland and LUKE by FMI and SYKE. The LUKE is responsible for LULUCF reporting GHG inventory for Finland by producing the estimates (emissions and removals) of all land use categories, and the text to the inventory report. The LUKE was subcontracted by FMI. That's why user engagement with the stakeholder was automatically achieved. Several online meetings and communication by emails were conducted with LUKE.

FMI, **SYKE** and **Luke** organized a user engagement event on the 12th of May 2021. This was a webinar for current proceedings in monitoring land use and land use change and reporting LULUCF in Finland and European level. The webinar was very successful in bringing together stakeholders working on the field of land use monitoring and GHG-inventory and fruitful discussion on current and future cooperation was had. Altogether 56 people attended the webinar.

ther In March/2020 **SRTI-BAS** organized a two-day (11-12 march) online workshop, entitled, "Copernicus program – access to products and services supporting/enhancing the inventory process of GHG emissions and removals form Land use, Land use change and Forestry (LULUCF)". The aim of the event was to present information on Action 2019-2-49 DEVELOPING SUPPORT FOR MONITORING AND REPORTING OF GHG EMISSIONS AND REMOVALS FROM LAND USE, LAND USE CHANGE AND FORESTRY and results from the work of the Bulgarian team (SRTI-BAS and CASTRA) within the project. Experts from SRTI-BAS, CASTRA, Ministry of environment and water(MOEW), Forest Research Institute(FRI-BAS), Climate, Atmosphere and Water Research Institute(CAWRI-BAS), National Statistical Institute, Bulgarian Development Bank, NGO "GEO Polymorphic Cloud", NGO " Eco horizons" participated in the workshop. Additionally, a discussion on "Perspectives and challenges in Copernicus data for improving environment monitoring and management" was held on the second day of the workshop.

Main issues discussed referred to possibilities for GHG emission inventory responsible institutions to use Copernicus data.

SRTI –BAS organizes another user engagement event after 20th of May.

On 18th March of 2021, **IGIK** organised Training Session on the Use of Copernicus Program Products (Training) as part of the FP CUP - DG CLIMA project. The training was held on-line via the MS Teams platform. It was an event that runs from 9:00 am to 2:30 pm. Following topics were covered,

- FP CUP project and DG Clima project – overview
- Introduction to the Copernicus Program, overview of the Sentinel mission
- Overview of LAND Copernicus vector layers – Corine, Urban Atlas, Riparian Zones
- Overview of LAND Copernicus raster layers – HRL: Imperviousness, Forest, Grassland, Water and Wetness
- Overview of LAND Copernicus global products + use of Copernicus data by Google Earth Engine
- CAMS

The training was attended by 28 people who are employees of the Institute of Environmental Protection of the Polish Academy of Sciences – IOŚ-PAN. Participants from:

KO – Krajowy Ośrodek Bilansowania i Zarządzania Emisjami

OR – Ośrodek Zrównoważonego Rozwoju

CC - Krajowy Ośrodek Zmian Klimatu

BA - Zakład Modelowania Atmosfery i Klimatu

BM - Zakład Ocen Środowiskowych, Ochrony Przyrody i Krajobrazu

BI - Zakład Chemii Środowiska i Oceny Ryzyka

BN-Stacja Kompleksowego Monitoringu Środowiska "Puszcza Borecka"

BJ - Zakład Ochrony Wód

BL - Centralne Laboratorium Analiz Środowiskowych - CentLab

KE - Zespół Inwentaryzacji i Raportowania Emisji (KOBiZE)

KB - Zespół Zarządzania Krajową Bazą (KOBiZE)

KZ - Zespół Programów Redukcji Emisji (KOBiZE)

CE- Zakład Niskoemisyjnego Transportu

In Spain, **existing gaps** were defined obtained through a comprehensive process of stakeholder engagement, Copernicus user uptake and participatory process made in close collaboration with the Spanish Inventory System (SEI) during the whole year 2020. There were celebrated a total of 16 meetings on which relevant stakeholders and different policy departments of MITERD that allowed defining and discussing a number of gaps related to the four main categories of land use and cover types included in LULUCF regulation (i.e. Forest land, Grasslands, Croplands and Wetlands). Prior to COVID restrictions, there were developed 3 face to face meetings with MITERD in Madrid, with the objective of presenting the project, setting the scene and start putting in common LULUCF and Copernicus objectives and future developments. There were also developed 2 more important meetings on which there was disseminated the project purpose in the context of Copernicus data and methods. The first corresponded to the SENSAGRI meeting developed in Madrid (see http://www.cdti.es/recursos/eventosCDTI/7654.pdf), with more than 100 accumulated attendees from administration and private companies of Spain and Europe. The second was developed at MITERD headquarters as the first step towards data organisation to identify information gaps for LULUCF reporting at the Spanish level. This meeting has been followed by 8 online workshops during 2020 in the base of COVID guidelines, not being able to develop as face-to-face meetings. These meetings were structured in working groups with MITERD corresponding to LULUCF typologies, i.e. Forests, Grasslands, Croplands and Wetlands and their objective was getting a complete understanding of LULUCF requirements, data available and information gaps existing in ITERD for being used in current monitoring and reporting tasks. Finally, after a continuous discussion undertaken during the whole year 2020, there was a final online meeting at the beginning of 2021 to present and explain the final list of data gaps that can be found here, https://lulucf.ihcantabria.com/encuesta/. Once set up, there have been developed during April 2021 two additional workshops with more than 40 participants each day corresponding to the Advisory research group of the project, that is composed by LULUCF inventory experts, land-use experts, scientists, and stakeholders. They were invited to fill the above survey will all identified gaps and to exchange experience and expertise to ensure that the best possible use can be made from geographic-explicit land use data from the Copernicus Land Monitoring Services and related data and services for fulfilling the defined gaps. The link to these Copernicus for LULUCf workshops can be found at the web of the Spanish partnership of the project, https://lulucf.ihcantabria.com/eventos/# and in the web of ApliCop, related to the Copernicus User Forum by allowing mapping the Spanish ecosystem of Copernicus intermediate and final users, https://aplicop.ihcantabria.es/eventos-en-espana/listado-de-eventos/.

User engagement in the **Czech Republic** done via organising workshops and trainings for end-users. They can be listed as follow

- It was organized two online meetings (24/9/2020 and 7/10/2020) with the most relevant stakeholders of LULUCF - participants from IFER - Institute of Forest Ecosystem Research, Ltd (responsible person for LULUCF reporting in the Czech Republic: Dr. Emil Cienciala).
- It has been organized several discussions on the current greenhouse gas inventory methods and the actual methodology of LULUCF detection in the Czech Republic with the members of the GEO/Copernicus National Secretariat (e.g. 13/7/2020, 28/1/2021).
- It was organized an online workshop with stakeholders of the land cover/land use databases (partners of Plan4All.eu) on 15/4/2021. It was presented concept and results of DG Clima FCUP project and discussed on using/combining spatial-based databases in the LULUCF reporting (e.g. Open Land Use Map)
- The final stakeholder workshop is planned for the middle of June 2021 with a cooperation of GEO/Copernicus National Secretariat

3.2. Test cases in selected NUTS

Each partner selected a study area in their country (*Table 3.1*). The average size of NUTS2 in each country is different shown in Table 3.x. We propose to choose NUTS2 with an area larger than 20000 km2. In the case of small NUTS2, two NUTS2 areas could be taken. Also, a part of big NUTS2 can be selected but it must comply with country rules of GHG reporting (usually the whole country has been reported).

Table 3.1: Characteristics of selected NUTS2

Country	Area of country km2	NUTS 1	NUTS 2	Aver. area of NUTS 2
Bulgaria	110 994	2	6	18499
Czechia	78 865	1	8	9858
Finland	338 440	-	-	-
Ireland	70 274	1	3	23425
Poland	312 679	7	16	18393
Spain	505 990	7	19	26631

In **Finland,** the test cases were performed for the whole country area (not a single NUTS2). This approach was chosen as the scale of the complementary atmospheric top-down approaches for monitoring LULUCF related emissions was more suitable for a large area. Also comparing the calculations of LULUCF emissions with current national methodology and with Copernicus data was considered more suitable for the country as a whole. All the necessary data was available for the whole country.

The criteria for choosing the NUTS2 in **Poland** was considering the Administration Division with the variation of land use types. The Podlaskie Voivodeship (province) has been chosen as one of the 16 NUTS2 Administration Division in Poland. It is a voivodeship in northeastern Poland, area 20 187,02 km². The province has variable vegetation in its area and density. The vast forests, some of which are the only ones in Europe character, contain a unique wealth of flora and fauna. The vegetation of the region is extremely diverse. The natural conditions of the region allow to develop vast areas of grasslands for cattle breeding. Also, in Podlaskie Voivodeship the large wetland is located with open water areas and high soil moisture. In several sites in Biebrza wetlands and surrounding grasslands, IGIK has carried out the measurements of Leaf Area Index and biomass as well as acquisition of S1, S2/S3 data.

The criteria for choosing the NUTS2 in the **Czech Republic** was the Administration Divisions with the dynamics and variation of land use types. The NUTS 2 Central Moravia (CZ06) and South East (CZ07) have been selected. The NUTS2 have variable land use types and heterogeneity from the point of view of physical and social geography. The forest has been influenced by rapid bark beetle attack with strong impact on the health of the forest ecosystems.

The test case for **Bulgaria** (land cover mapping performed by SRTI-BAS) selected NUTS is South eastern (Yugoiztochen) region. However, the research was performed for the whole country area. This approach was chosen as the national greenhouse gases inventory report provides data at country scale only. Thus, to compare the land use areas calculated within our test case with those derived with current national methodology it was necessary to map the country as a whole. In addition to that, Bulgaria is characterised with diversity of geographic conditions (climate, altitude, relief, vegetation, land use/land cover, etc.) that could not be represented by any single NUTS2. Thus, a national scale test case was considered more representative and useful.

In **Spain**, following the process of participatory learning explained above, we are currently collecting and defining solutions for the gaps identified. Among all of them, we will select 6 success stories that will be applied and carefully presented in selected NUTS. Two of them will correspond to the Autonomous Region of Cantabria, located in northern Spain and one to the Province of Malaga, from the Andalusían region of the South of the country. 4 more are still to be defined awaiting reviewing all the responses of the Advisory research group of the project to the submitted formulary with the gaps. Deadline for this task is 2021 May 14th.

3.3. Proposed solutions and methodologies for use in national reporting

3.3.1. Bulgaria (SRTI-BAS & CASTRA)

SRTI-BAS tested two approaches for land cover/land use mapping: the first used satellite image classification, while the second was based mostly on Copernicus HRL data. The methodology and results of both approaches will be presented separately.

The land cover maps prepared using the first approach included the following classes: Cropland, Orchards and vineyards, Grasslands, Shrublands, Broad-leaved Forest, Coniferous forest, Water, and other land cover.
Settlements were masked out before the classification. Reference data about the land cover class in selected points were collected in terms of visual interpretation of high-resolution imagery from Google Earth. Sample was constructed in which each point has the same land cover over the entire period 2012-2018. Thus, the same reference points were used for all classification years. The 4000 points were split at random in two datasets used for training of the classifier (70% of points) and validation of the classifications (30% of points) respectively. Different type of satellite imagery was used for the classifications in different periods (Landsat 7 in 2012, Landsat

8 in 2013-2015, and Sentinel-1/2 in 2016-2018). The satellite data are available in Google Earth Engine (GEE; Gorelick et al., 2017). GEE was used for the image pre-processing and random forest classification.
Figure 3.2 shows the land cover map for 2018 derived by classification of Sentinel-1 and -2 imagery. The accuracy statistics are graphically summarized in *Figure 3.3*. The OA varies from 66% in 2013 to 81% in 2018. Clear tendency of increasing the OA with years is observed. More importantly, the maps produced by using Sentinel data are more accurate (79% on average) than those using Landsat (69% on average).

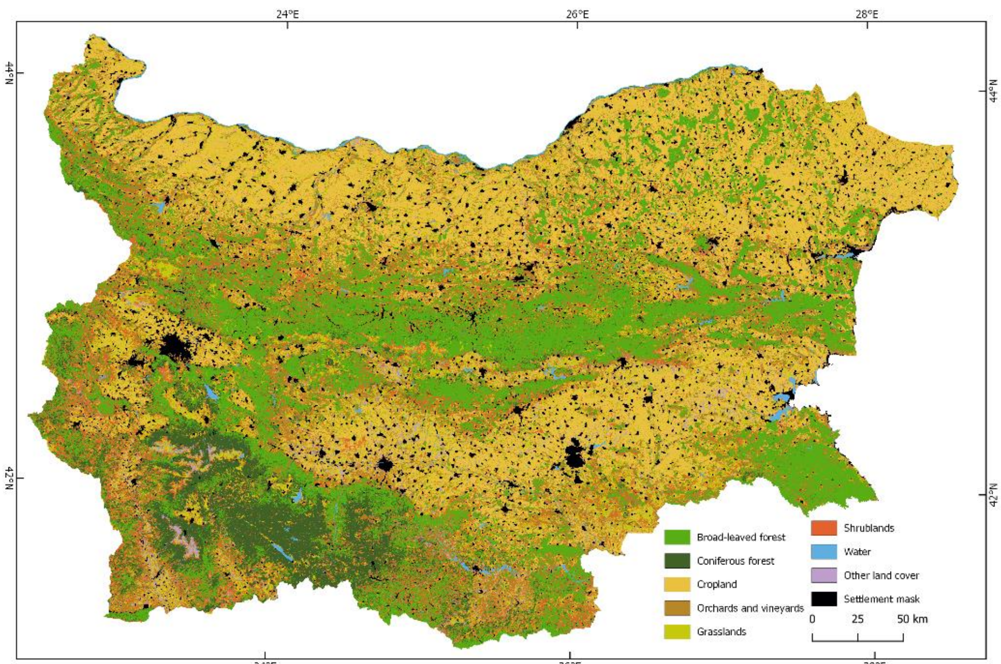


Figure 3.2: Land cover map of Bulgaria for 2018 derived from Sentinel-1/2 imagery.

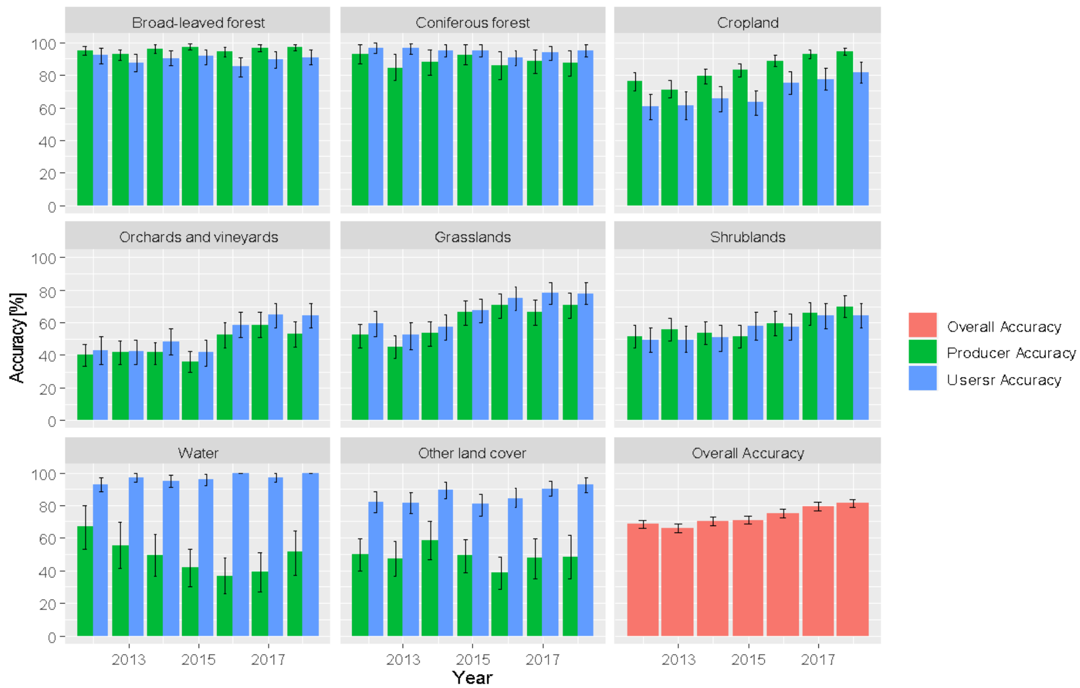


Figure 3.3: Accuracy statistics and 95% confidence intervals.

The land cover maps for 2015 and 2018 prepared using the second approach included the following classes: Settlements, Agricultural land (both annual and perineal crops), Grasslands, Broad-leaved forest, Coniferous forest, Wetlands, and other land cover. The main data source to map those classes were the Copernicus HRLs and Corine Land cover (CLC) 2018. *Figure 3.4* shows which dataset/datasets is used to map each class as well as the priority rule in case of overlap between datasets.

Dataset / Subset	Mapped class
Corine Land Cover 2018 / Class 100 Artificial surfaces + HRL Imperviousness Density (IMD) / Cover 1-100%	Settlements
HRL Tree cover density (TCD) / Cover 10-100% + HRL Dominant leaf type (DLT) / Broad-leaved forest	Broad-leaved forest
HRL Tree cover density (TCD) / Cover 10-100% + HRL Dominant leaf type (DLT) / Coniferous forest	Coniferous forest
HRL Water and wetness (WAW) / Permanent water, Temporary water, Permanently wet	Wetlands
HRL Grassland (GRA)	Grasslands

Figure 3.4: Copernicus datasets (CLC and HRL) to map some land cover classes in Bulgaria. The higher the dataset/class in the table, the higher is its priority in case of overlap with the other datasets.

Combining the above-mentioned existing Copernicus datasets was not sufficient to produce wall-to-wall map of the country. Land cover in areas not covered by the datasets had to be mapped by other means. First, these areas were divided into two hypsometric belts, i.e. above and below 1400 m a.s.l respectively and their predominant land cover analysed using Google Earth imagery. In the still unmapped areas above 1400 m a.s.l the land cover was represented mostly by the classes Coniferous forest (e.g. Pinus mugo), Grasslands, and Other land cover (e.g. bare rocks). In the areas below 1400 m a.s.l the land cover was represented mostly by the classes Agricultural land, Grasslands, and other land cover (e.g. bare rocks, beaches). Though other land cover classes may also be present it was assumed that this was not the case to simplify the further analysis. The areas with missing data were filled in using Sentinel-2 and Landsat 8 data by two means. First the Other land cover class was mapped assuming that it represents the areas where the value of the Normalised Difference Vegetation Index (NDVI) is below a certain threshold. Then random forest classification of Sentinel-2 (2018) and Landsat 8 (2015) imagery was used to map Coniferous forest, Grasslands, and Agricultural land. The resulting land cover map for 2018 is shown in *Figure 3.5*. The accuracy assessment of the maps was performed using reference data from Google Earth. The results are shown in *Figure 3.6*. *Figure 3.7* shows the comparison of the land cover areas in Bulgaria derived by the National Greenhouse Gasses Inventory Report and the land cover maps.

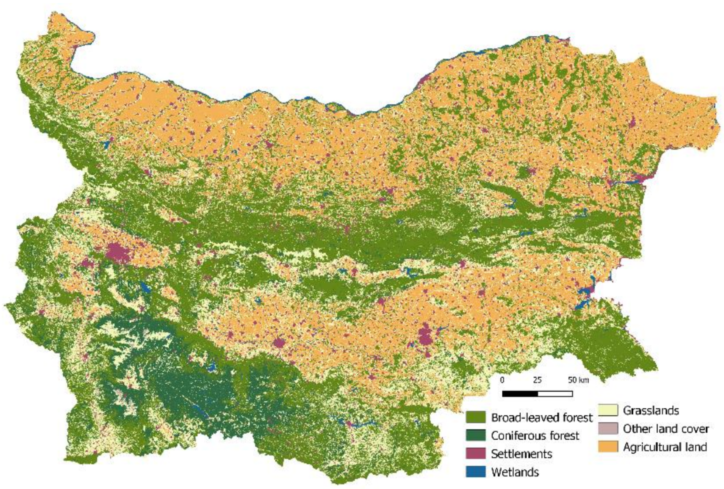


Figure 3.5: Land cover map of Bulgaria for 2018 derived by existing Copernicus datasets (CLC and HRLs) complemented with land cover information extracted from Sentinel-2 by image classification and vegetation index thresholding.

2015	Area [km²]		Producers' Accuracy [%]		Users' Accuracy [%]	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Broad-leaved forest	34404	2863	90	5	85	6
Coniferous forest	11143	1732	79	11	91	7
Settlements	5641	970	87	13	85	8
Wetlands	1462	816	58	33	96	4
Grasslands	19292	3471	71	9	58	11
Other land cover	166	151	54	49	44	11
Agricultural land	38887	3029	81	5	91	5
Overall Accuracy 81,9 ± 3,5 %						

2018	Area [km²]		Producers' Accuracy [%]		Users' Accuracy [%]	
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Broad-leaved forest	37232	2601	93	4	89	5
Coniferous forest	10204	1406	84	11	95	5
Settlements	7502	1707	65	14	85	8
Wetlands	1364	619	72	32	93	6
Grasslands	19469	3273	81	8	63	10
Other land cover	433	594	30	41	53	11
Agricultural land	34786	2492	85	5	96	4
Overall Accuracy 85,2 ± 3,3 %						

Figure 3.6: Accuracy statistics and per-class area estimates.

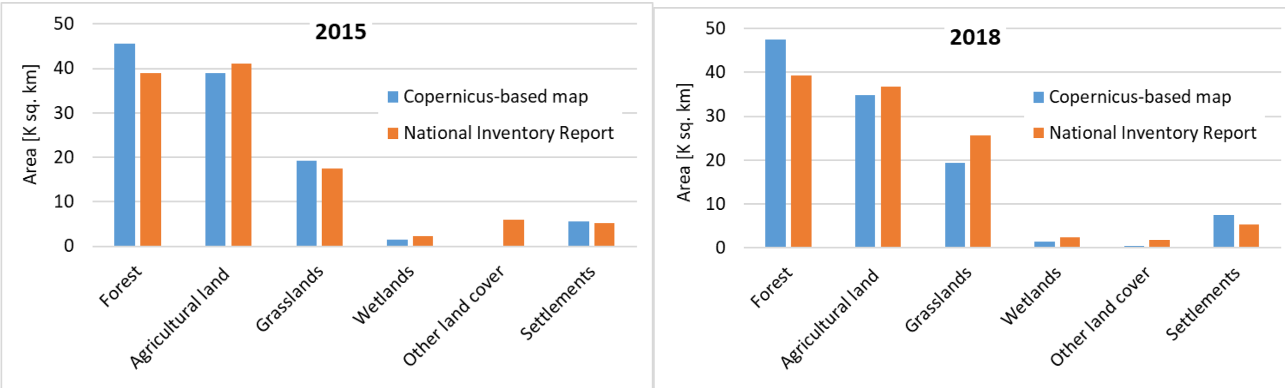


Figure 3.7: Comparison of the land cover areas in Bulgaria derived by the National Greenhouse Gasses Inventory Report and the land cover maps based on Copernicus data prepared as part of this test case.

In addition to the above works, CASTRA also focused on investigating the possibilities to develop a general methodology aiming establishing a single unified LULUCF layer that could track in a spatially explicit way the land use changes over time. In this approach, CLMS and others EO systems can be used to supplement, enhance and/or verify the national datasets used by the reporting authorities and stakeholders. This is expected to ensure complete land use representation (all LU categories will be included in the layer). The single layer can be achieved by combining all spatial data available in the country (LPIS, FMP, cadastre etc.) for a certain piece of land, to analyse the gaps and possibilities to combine these datasets with pan-EU remote sensing data and achieve a complete and accurate spatial dataset for LULUCF reporting. Specific attention towards independent verification or expanding the temporal coverage of the land use dynamics could be made for FL and GL by considering the possibilities of HRLs that CLMS provides, based on the results described above. The methodology requires the definition of a workflow and tools to automate the processing of the available spatial data for the country, for the Integration and harmonization of the available datasets, definition, and semantic correspondence with the land use categories for LULUCF., selection of proper remote sensing CLMS/EO datasets to complement the land representation., validation and test application. The integration of different spatial datasets into one single layer is linked with many technical issues related to geoprocessing of the information to harmonize the data which is e.g. available with different temporal and spatial scales, and which data is usually generated in different formats and coordinate systems. This is found to require developing (as a test case here) the unified LULUCF layer at NUTS 3 level instead of NUTS 2 because of the workload expected. This restriction is further analysed.

A specific concept to resolve the difficulties of creating a unified LULUCF layer out of several different datasets, is the possibility to develop novel instruments and methods for geospatial datasets combination, using semantic technologies -based database manipulation. This approach potentially allows for an 'effort savvy' and easy maintenance of the incoming information even when it is necessary to change some of the original data specifications. This is because the semantic- based approach do not require strict design of a rigid database parameter exchange scheme, which in turns requires redesign of the entire system, as in the 'classical' relational databases. The semantics-based data extraction methods make it possible to create flexible interfaces between the original individual datasets, and thus make it possible to retrieve and use portions of data as needed without having to physically transfer the original database to a new location (e.g., end user), but only requires describing the information about the data and use it in the new layer/dataset (see *Figure 3.8*).

In addition to the ease of maintenance and distributed management, this approach of using semantics based data retrieval technologies can also benefit from a large set of already existing and publicly available connected data on global scale, for example open sources at http://www.linkedeodata.eu/Linked_Geospatial_Data, http://www.opengeospatial.org/, http://www.opengeospatial.org/standards/sos - OGC Sensor Observation Service (SOS) Standard Version 2.0 Adopted, http://www.earthobservations.org/geoss_dsp.shtml - Group on earth observation

Such other global databases, together with CLMS/Copernicus databases and using the existing 'domestic' LULUCF reporting datasets, can form a solid (and flexible) pool of Geospatial data sources (as illustrated on the *Figure 3.9*), used for creating a unified LULUCF layer as discussed above, applying semantics-based data retrieval technologies.

Implementing the above framework for combining different geospatial datasets seems very appropriate to establish a flexible approach for creating maximally enhanced local and pan-European databases for LULUCF reporting, in which CLMS/Copernicus data can serve as a backbone. It will be needed to investigate further if more specific requirements for development of new Copernicus products better suited for semantics-based data manipulation, will be needed.

On-going and further works are planned for a software tool implementation of the above approach of combining different geospatial datasets towards creating an unified layer aiming land use time-tracking for the purpose of LULUCF emissions/removals estimation and monitoring.

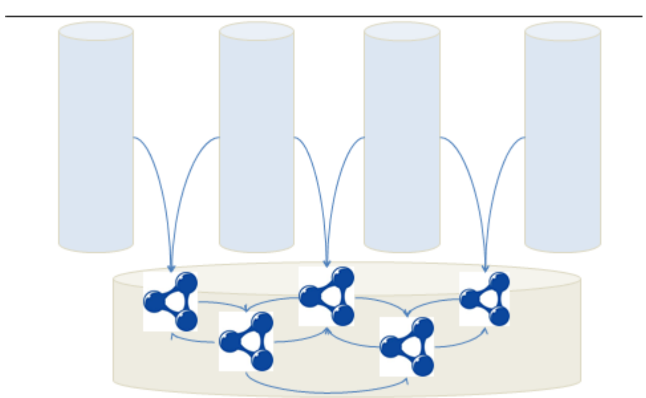


Figure 3.8: Illustration of the concept of 'extracting' data elements needed to form a custom new data layer from a set of existing datasets. Semantics-based methods for data manipulation seem most appropriate.

Figure 3.9: Illustration for creating a pool of different Geospatial Databases (non-standardized, remote, heterogeneous), which can be 'linked' through semantics-based methods for data manipulation for the purpose of creating a new custom data layer.

3.3.2. Czech Republic (CUNI)

LULUCF mapping in two NUTS2 regions in Czech Republic (Southeast and Central Moravia) has been carried out for years 2017, 2018 and 2019. The Copernicus Sentinel-2 data have been used and Random forest classifier in the open access platform of Google Earth Engine applied and tested. The results of classification have been validated/assessed. For this purpose, control points have been collected using in-situ data as well. Classification has been carried out following the LULUCF regulations: with 10 meters minimal mapping unit, with accuracy higher than 80 % and annually (years 2017, 2018, 2019).

Training polygons were created by 2 methods. The first method is the semi-automatic creation of training polygons from the CORINE Land Cover layer for the year 2018. From the CLC2018 polygons, core areas were created first using the Buffer function - 200 m (ArcMap). Inside these areas, training polygons in the shape of a circle with a diameter of 80 m were then randomly generated. Then were checked whether these polygons are on the declared Land Cover class using aerial photographs and topographic maps (CIR, ZM10, Ortofoto WMS by ČÚZK; Aerial photographs / satellite images by Mapy.cz, Google Earth Pro and Planet.com - PlanetScope). In this way, 232 training polygons were generated.

The first method of creating training polygons did not include all types of surfaces that are present in the area of interest. For example, no training polygon was created on the territory of a photovoltaic power plant and no generated training polygon could be considered as the other land. Training polygons for these surfaces were manually added in a number of 34 training polygons (*Figure 3.10*).

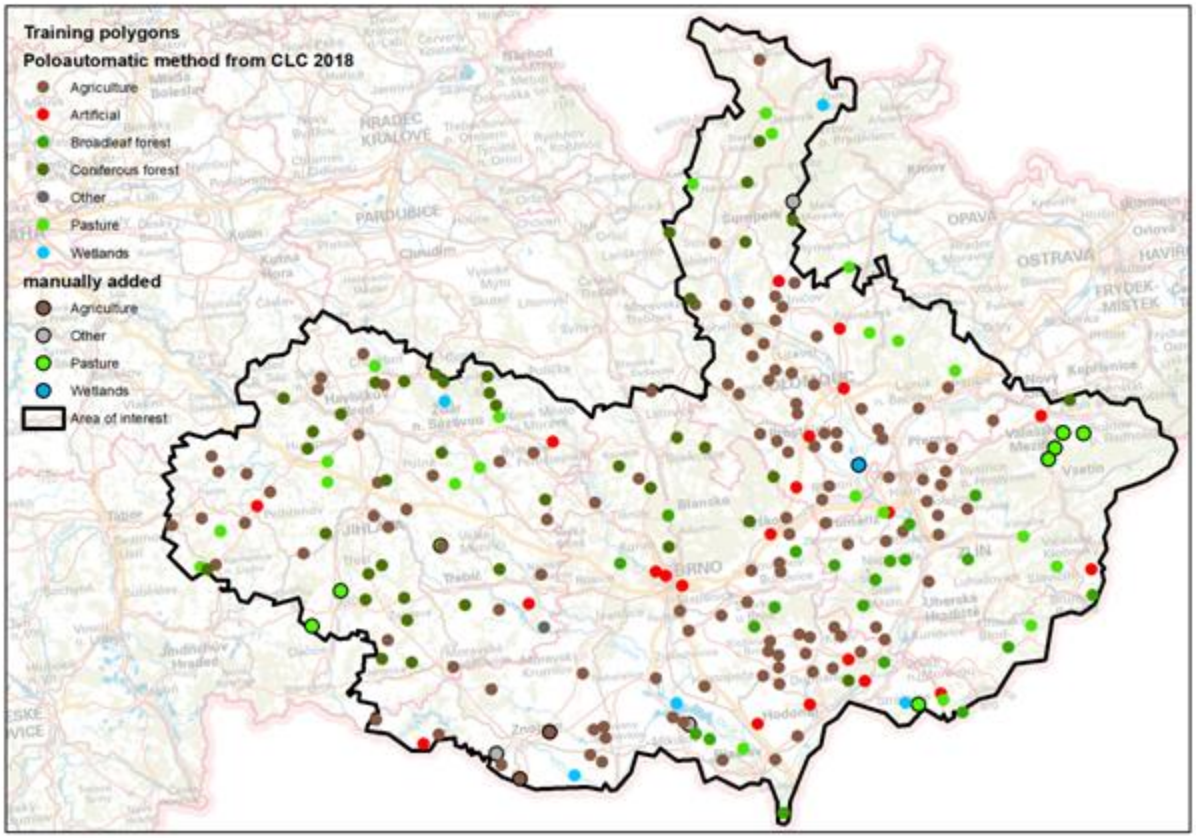


Figure 3.10: Training polygon locations.

Mosaics for classication were created from cloudless pixels taken during the May-July period in each analysed year. For each pixel, a 40% percentile of the cloud-less values was determined as the final value. This procedure was used due to unmasked clouds, which would increase the median or average value up. This procedure was chosen for the values of all S-2 bands with a resolution at least 20 m (without band B8) and the NDVI index (calculated from bands B4 and B8). The mosaic also includes a band representing the variance of NDVI values in the period from May to October. This band helps to distinguish surfaces as buildings (small variance) from surfaces with dynamically changing NDVI (arable land).

Random Forest classifier (RF) was chosen as most accurate and was used for classifications (chosen by Testing Control Points). The samples were tested. After testing of spectral characteristics/classification inputs, main categories were divided into subcategories with cluster analyses before classification for each year and after classification again was aggregated.

The classes were divided like this because the LULUCF class can include a variety of surfaces (e.g., agriculture includes wheat fields, pea fields, vineyards etc.) and visually it is not possible to reliably divide individual training polygons into subcategories because more detailed information about training polygons is missing.

The majority filter was applied after classification (low-pass filters: https://developers.google.com/earth-engine/image_convolutions).
Final output was created with the applied low-pass filter (majority filter for reduction of pixels which share an edge with one or zero neighbor from the same class) in 10 m resolution and in the UTM 33N projection (EPSG 32633). 1980 combinations of parameters (Number of Trees, Variables per Split, Bag Fraction) were tested for each year and the best combination were used for classification (selected by OA using Testing control points).

Results were validated with use of aerial photos and topographic maps (CIR, ZM10, Ortofoto WMS by ČÚZK; Aerial photographs / satellite images by Mapy.cz, Google Earth Pro and Planet.com – PlanetScope).
Based on the Test control-points, the classification for 2017 has an accuracy of 83.78% (the kappa index value is 78.6). *Table 3.2* shows validation of individual classes. The biggest problem is in distinguishing the Other land class from Settlements and classification of grassland like arable land.

Table 3.2: Validation of individual classes.

	Settlements	Agriculture	Forestland – broadleaf	Forestland – coniferous	Grassland	Wetlands	Other land
Settlements	**38**	3	0	0	1	0	0
Agriculture	5	**265**	0	1	9	4	0
Forestland – broadleaf	1	5	**93**	12	5	0	0
Forestland – coniferous	0	2	6	**106**	1	1	0
Grassland	6	31	1	1	**101**	1	0
Wetlands	0	0	0	0	1	**8**	0
Other land	23	0	0	0	0	0	**9**

The area of individual LULUCF classes for 2017 can be seen in the table below. Arable land is dominant in the area of interest.

LULUCF class	Settlements	Agriculture	Forestland	Grassland	Wetlands	Other land
2017	8.65	41.35	33	15.14	1.77	0.09

Classified area of interest for 2017 is shown in *Figure 3.11*.

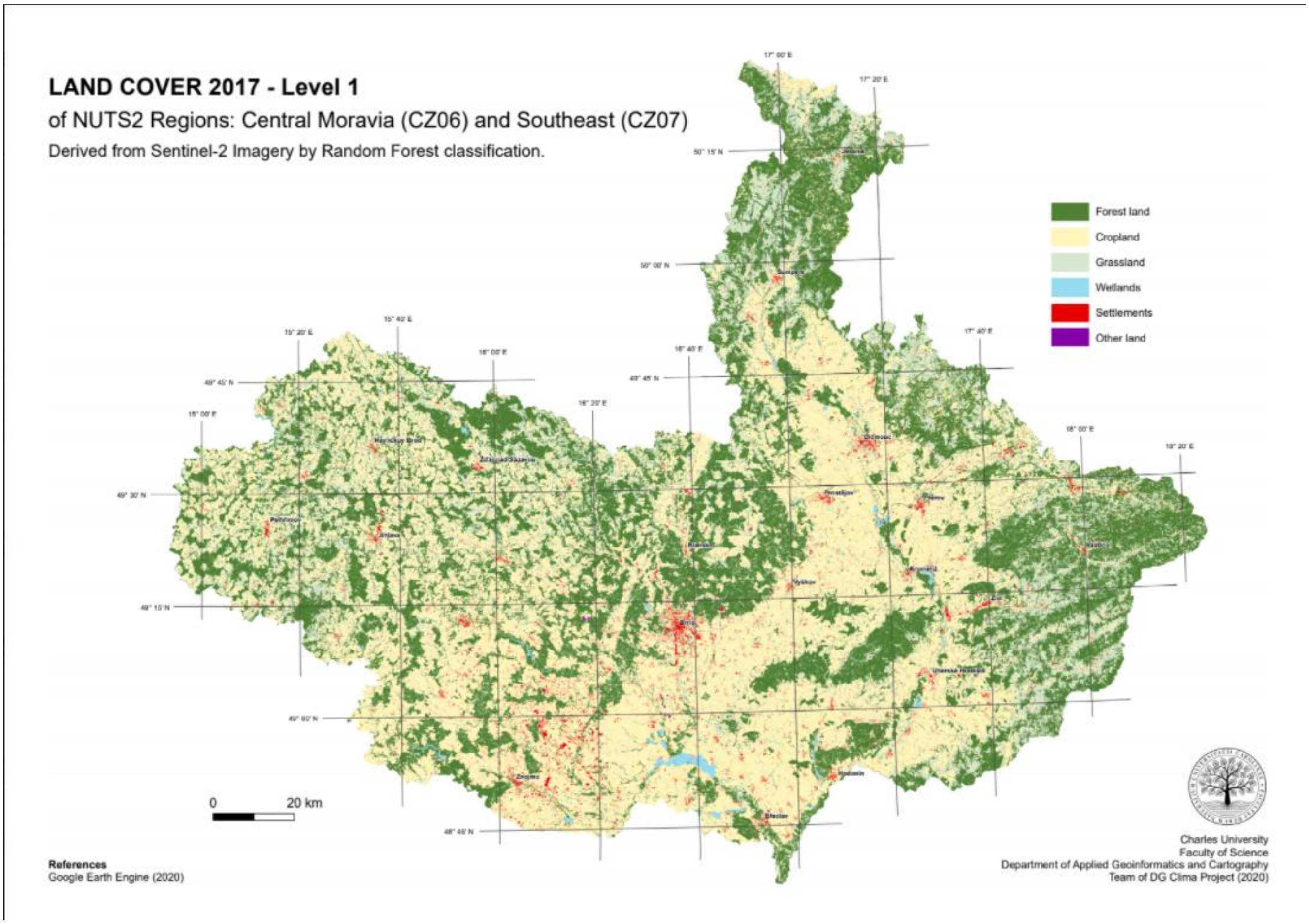


Figure 3.11: Classified area of interest in 2017.

3.3.3. Finland (FMI (Luke & INRAE) & SYKE)

Comparison of Finnish GHG inventory land use and land-use change area data to Copernicus HR CLC data

Foreseen potential of Copernicus products in LULUCF reporting in Finland were tested by using the CORINE Land Cover datasets from 2000, 2006, 2012 and 2018. These national HRL datasets provide information on Finnish land cover and land use, and its changes. The CORINE 2018 data was produced as a part of the EU Copernicus Land project (SYKE 2020). CORINE land-cover classes were classified into the GHGI LULUCF land categories together with SYKE. The CORINE data has information on mineral and organic soils. In a case where the information on soil type was missing, the Finnish soil database was used. The advantage of Finnish CORINE data is that it has a high resolution of 20 m pixel size or 25 m for older layers. Therefore, it represents what future European level Copernicus data could be.

Land-use areas were produced first for the years 2000 to 2018 (*Figure 3.12*). We assumed that, for example, the CLC 2018 represents year 2018 and CLC 2000 represents year 2000. For each land-use category, the total area was derived from 2018 CORINE CLC layer, and land-use changes from CORINE change layers. Using these changes from 2018 backwards, time series 2000-2018 for each land-use category was constructed. CORINE CLC resulted in larger forest land area than GHGI. It can be noted that part of woody land in GHGI is under WL and OL. That is in case they do not fulfil the minimum thresholds for forest land definition. OL areas in GHGI are more frequent especially in Southern archipelago with rocky cliffs and in Lapland high lands, where climatic conditions also limit land productivity. Larger wetland areas in GHGI appear throughout the country compared to CLC. Cropland in the GHGI includes field parcels that are not in Land Parcel Identification System (LPIS) data related to subsidies by EU. Therefore, in the GHGI cropland area was higher than in CLC data. Cropland not included in LPIS data in CLC were included in grasslands. However, they included large areas of croplands according to NFI field survey. This indicates less intensive cultivation. This type of fields explained the most part of difference in areas between CLC and GHG inventory on croplands.

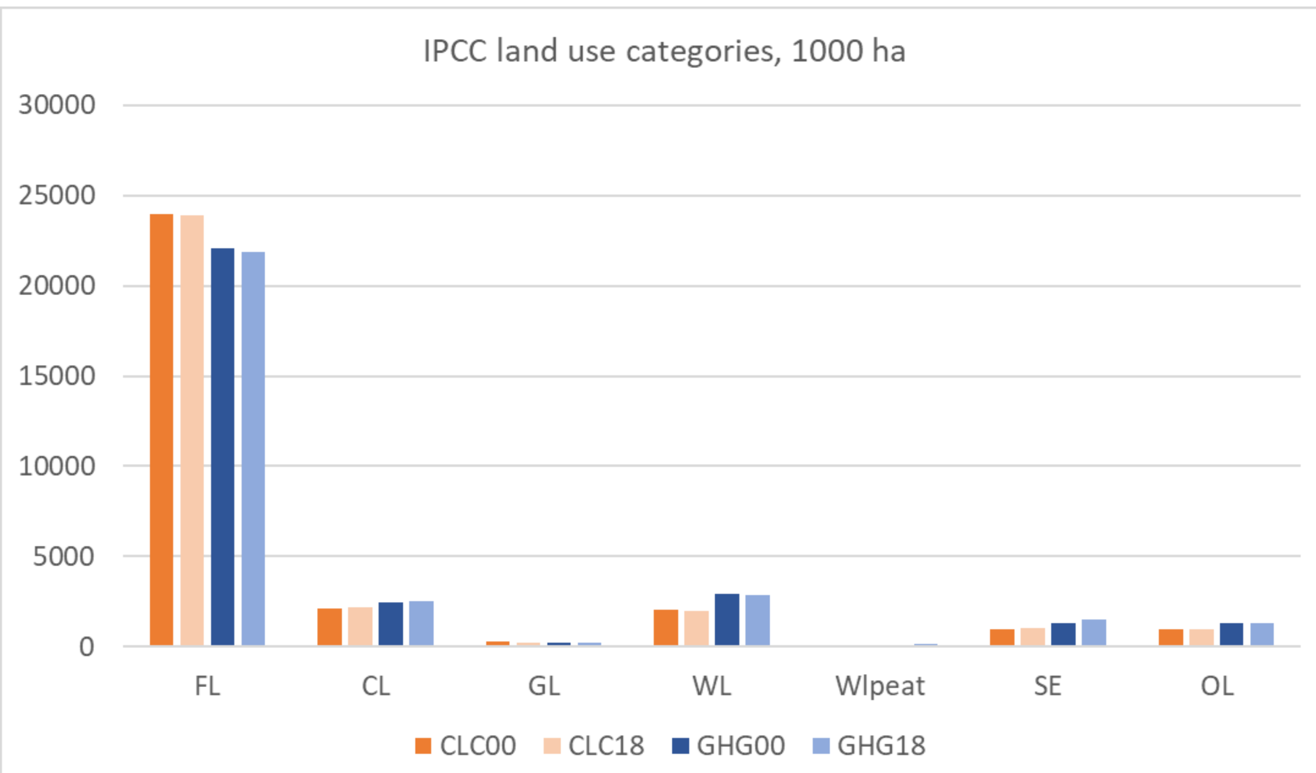


Figure 3.12: CLC based areas and GHGI areas by land use categories in years 2000 and 2018. FL=forest land, CL=cropland, GL=grassland, WL=wetlands, Wlpeat=wetlands/peat extraction, SE=settlement, OL=other land.

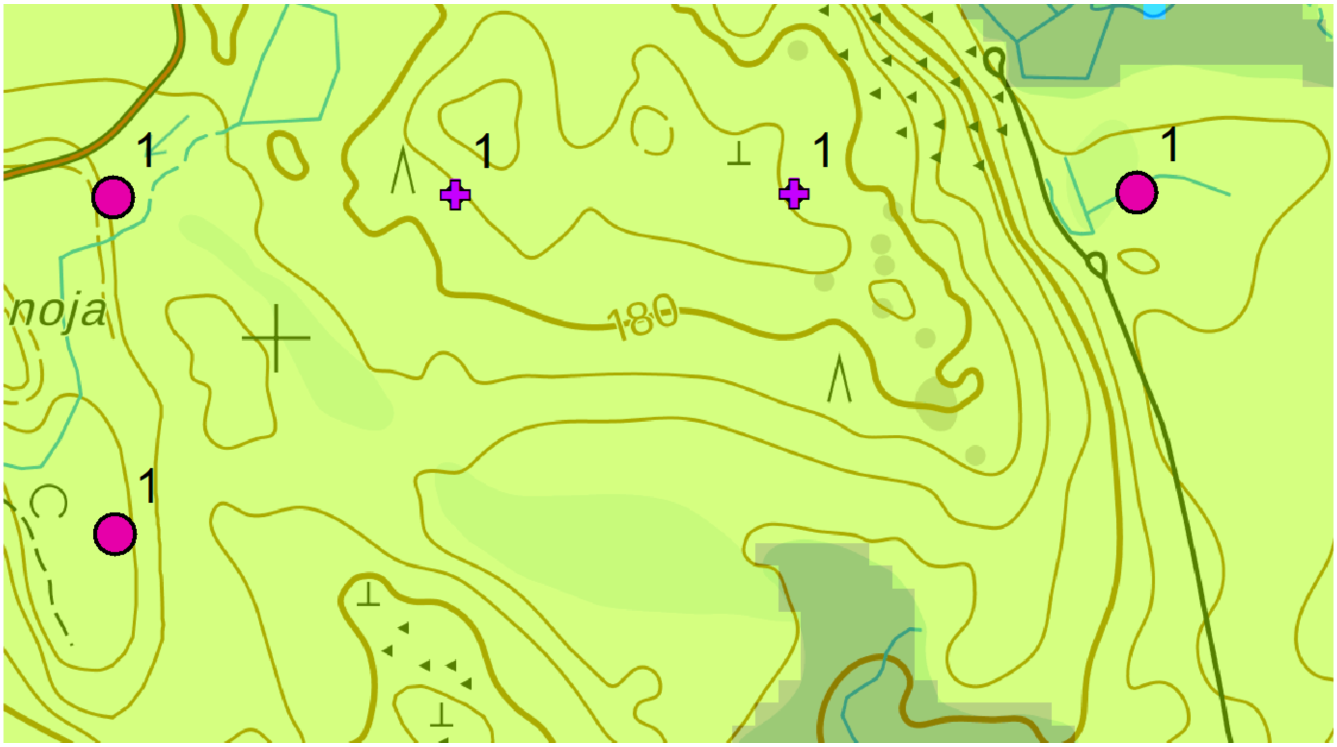


Figure 3.13: NFI plots on forest land, where circles indicate organic soils and crosses indicate mineral soils. Dark raster is CLC organic soil. Especially areas with thin peat layer are often mineral land according to CLC. Transparent NLS map layer is in the front.

Peatland soils in CORINE CLC on wetlands and forest land were derived from available numerical maps provided by national land survey of Finland. However, in NFI field survey peatlands were defined by vegetation and are therefore larger than according to map layers (*Figure 3.13*). On croplands and grasslands, CORINE CLC did not separate organic soils from mineral soils. Therefore, soil database was used as an additional data source. It was used also in GHGI for croplands and partly for grasslands.

It was assumed that annual change areas before 2000 were the same as in 2000 based on 2006 CORINE change layer, which is not correct. The GHG inventory employs National Forest Inventory (NFI) data, which are available from 1970. For the older years, hybrid data could be the solution to produce more accurate estimates (combination of CLC, NFI, and other sources). Proportions of other sub-categories, like site type for drained forest lands, were derived from NFI and applied to Corine CLC data.

Land use matrix between GHGI and Copernicus HR CLC data is presented in Table 4. GHG areas are derived from NFI plots. Number of NFI plots does not correspond directly to the area due to differences in sampling, which is more intensive in the Southern Finland and sparser in the North. Hence, CLC forest land had 6% higher number of field plots and areas were over 9% higher than in GHG inventory. The reason is that the CLC classification is based on land cover. In the GHGI, especially definition of forest land has features of both land-cover and land-use classifications. Larger areas of woody lands are settlements in the GHGI but in the CLC data they are forests. Reclassification on CLC classes to IPCC categories were done by best fit method and expert judgement for each class. However, number of CLC classes with lower forest cover are non-forest areas in GHGI in many cases. In

Northern Finland, large areas of forestry land are below the threshold of forest definition in GHGI, and thus not classified as forest land.

On croplands in GHGI data, only 85% of the NFI plots were classified to cropland in CLC. On the other hand, 94% of CLC cropland plots were included in GHGI cropland. Grassland areas in GHGI and CLC were close to each other, however they included different land parcels. GHGI grassland includes areas with tree cover (abandoned fields with natural forest expansion) whereas these lands tend to be classified into forest land in CLC. Almost half of the GHGI category 'other land' is classified into forest land in CLC data. In the case of wetlands the results are slightly better.

Deforestation area shows declining trend both in CLC and GHGI data (*Figure 3.14*). Years 1990 to 2000 are not comparable due to the method applied for those years (see above). Overall, deforestation areas in CLC data are lower than in GHGI. That is due to small changes that are not recognised in CLC in the same way as in NFI, which considers land use as well alongside land cover.

Table 3.3: Land use matrix between GHG inventory and CLC data in 2018, values extracted for NFI plots. FL=forest land, CL=cropland, GL=grassland, WL=wetlands, Wlpeat=wetlands/peat extraction, SE=settlement, OL=other land.

	CLC							
GHG inventory	FL	CL	GL	WL	SE	OL	Wlpeat	Number of NFI plots
FL	47963	128	48	338	362	78	12	48929
CL	407	5357	366	12	188	0	3	6333
GL	278	130	119	10	27	0	1	565
WL	1280	2	5	2499	1	14	3	3804
SE	1584	97	28	62	2099	11	6	3887
OL	328	2	6	46	3	378	0	763
Wlpeat	44	6	1	3	4	0	203	261
Number of NFI plots	51884	5722	573	2970	2684	481	228	64542

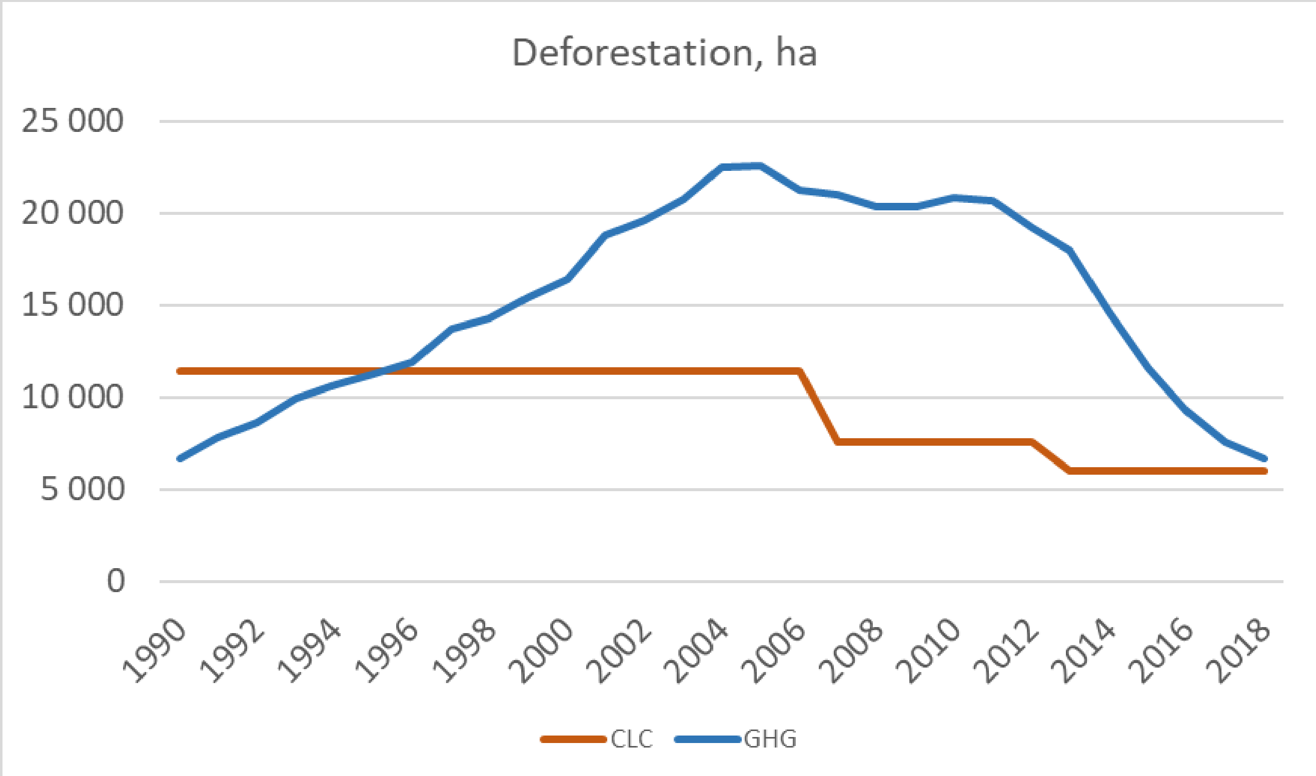


Figure 3.14: Deforestation areas in GHG inventory and based on the CLC data.

GHGI data and Copernicus HR CLC area-based estimates

Emissions and removals in GHGI reporting categories were calculated using the areas based on the HR CLC data. Presented results are net emissions or net removals, which consists of carbon stock changes of living biomass, soil, litter and dead wood pools, methane (CH_4) and nitrous oxide (N_2O) emissions from drainage of organic (peat) soils in forest land and wetlands, direct and indirect N_2O emissions from N mineralisation associated with loss of soil carbon due to land-use change on mineral soils, emissions from wildfires on forest land and grassland (CO_2, CH_4, N_2O), emissions from prescribed burning on forest land (CH_4, N_2O), and N_2O emissions from N fertilization on forest land. The harvested wood products (HWP) carbon pool was not included in the comparisons. Estimation of carbon stock changes in the HWP pool is based on the statistics on forest industry's production volumes of different product types. The Copernicus data do not provide input data to that kind of approach. Only emissions that are reported in the LULUCF sector under the Convention are included. Some emissions from soil are included in Agriculture sector, like N_2O emissions from organic soils and emissions from inorganic and organic fertilizers applied to soils on cropland and grassland.

The GHGI estimation procedures were carried out employing the HR CLC based areas. In this study, only the effects of different area source were examined. For forest land, to which the gain-loss method is applied, the gain was a product of mean biomass increment (kg/ha) derived from the GHGI data and the new area estimate. If the HR CLC land category would have been used on sample plot level, there would be non-forest sample plots included in forest land (see *Table 3.3*). To avoid effects of misclassification of sample plots, the GHGI mean values and emission factors were used. The loss of biomass is derived from the drain of growing stock (drain is due to harvest and natural mortality). The total annual drain is from statistics and is allocated to three parts: drain of total forest land, drain of afforested lands, and drain of deforested lands. The fourth part is drain of forest land remaining forest land, and it is calculated by subtracting the drains of afforested and deforested lands from the total drain.

The FI HR CLC area produced slightly higher net sink for total forest land and notably lower net emissions for cropland than the GHGI (*Figure 3.15*). In grassland and wetlands there is differences in levels, but the trends are similar. Category Lands converted to settlements shows the difficulty to detect land transitions with HR CLC data. The area in this category is in the GHGI twice the size of the changes detected from HR CLC data.

Figure 3.15: Net removals or net emissions in categories Forest land, Cropland, Grassland, Wetlands and Lands converted to settlement.

Regulation (EU) 2018/841 land accounting categories

From 2021 the LULUCF sector is included in the EU's climate and energy framework 2030. The Regulation sets out the commitments to the LULUCF sector so that the sector contributes to the Union's green-house gas reduction target and achievement of the Paris Agreement. The commitment for a Member State is to ensure that the sector's emissions shall not exceed the removals in periods 2021-2025 and 2026-2030. To follow the fulfilment of commitment, the emissions and removals of the LULUCF sector will be reported to the EU in six land accounting categories. The six land accounting categories are:

- Afforested land (AL)
- Deforested land (DL)
- Managed forest land (MFL)
- Managed cropland (MCL)
- Managed grassland (MGL)
- Managed wetlands (MWL).

Relation between the IPCC land use classification and EU LULUCF ac-counting categories

The IPCC land use categories and land-use change sub-categories (Statistics Finland 2020, IPCC 2006) were allocated into the EU's land accounting categories (Regulation (EU) 2018/841) (Table 5). This reallocation of lands is practicable, whereas some modification to calculation and output procedures are needed to allocate emissions and removals into the right accounting category. However, required changes are rather small. The accounting categories can be derived regardless of which data sources are used (e.g. Copernicus data), provided that all information needed for IPCC classification and emission estimation is either contained in that data or can be reliably complemented with information from other sources.

The six IPCC land-use categories and subcategories, that is land-use conversions between the six categories, are listed in six accounting categories according to the Regulation (Table 5). Settlements and other land are accounted for only in the case when changes to and from forest land, cropland, grassland, or wetlands occur. These cases are indicated in Table 5 in column 'NOT'. In practice, there are also transitions that hardly ever occur in Finland assuming the land classification remains as it is now. For example, conversions from other land to cropland have not been recorded, because other lands are barren lands and thus not suitable for cultivation.

GHGI data and Copernicus HR CLC area-based estimates in accounting categories

GHGI data and Copernicus HR CLC area-based emissions and removals were allocated according to the Regulation (EU) 2018/841. All emissions are not allocated exactly to the right accounting categories, but it does not hinder the comparison. For example, all CH_4 and N_2O emissions from drainage are included in the Managed forest land category in both cases.

A significant difference is found in Deforested land category (*Figure 3.16*). This indicates the difficulty to catch land-use changes related to large forest land area and rare and small size land transition areas, though in Afforested land category the results are more in line with each other. In MFL category, the HR CLC area-data based net sink follows nicely the GHGI based net sink of managed forest land (*Figure 3.16*), at the most providing about 1 Mt CO_2 eq. larger sink than GHGI data. Managed cropland is the only category in which the HR CLC area-based emissions are higher than GHGI based. MCL, MGL and MWL accounts are most unlike compared to GHGI results. The reason is, that these accounting categories include emission also from other land uses, and emissions of these land categories are included in Deforested land category (*Table 3.4*).

Figure 3.16: GHGI data and FI HR CLC based estimates according to the Regulation (EU) 2018/841 land accounting categories.

Table 3.4: IPCC land-use categories in the EU's LULUCF accounting categories. Convention reporting refers to the GHGI as reported to the UNFCCC. A=afforested land, D=deforested land, MCL=managed cropland, MGL=managed grassland, MWL=managed wetland, MFL=managed forest land, NOT=not ac-counted under the Regulation (EU) 2018/841.

Convention reporting			Regulation (EU) 2018/ 841 Accounting categories						
IPCC land use categories		From	AL	DL	MCL	MGL	MWL	MFL	NOT
Forest land	Remaining							X	
	Lands to FL	CL	X						
		GL	X						
		WL	X						
		SE	X						
		OL	X						
Cropland	Remaining				X				
	Lands to CL	FL		X					
		GL			X				
		WL			X				
		SE			X				
		OL			X				
Grassland	Remaining					X			
	Lands to GL	FL		X					
		CL				X			
		WL				X			
		SE				X			
		OL				X			
Wetlands	Remaining						X		
	Lands to WL	FL		X					
		CL			X				
		GL				X			
		SE					X		
		OL					X		
Settlements	Remaining								X
	Lands to SE	FL		X					
		CL			X				
		GL				X			
		WL					X		
		OL							X
Other land	Remaining								X
	Lands to OL	FL		X					
		CL			X				
		GL				X			
		WL					X		
		SE							X
Harvested wood products			X					X	

Geographically explicit land-use conversion data

The Regulation (EU) 2018/841 prescribes to use for monitoring and reporting in the LULUCF sector Approach 3: Geographically-explicit land-use conversion data in accordance with the 2006 IPCC Guide-lines for National Greenhouse Gas Inventories. The Regulation on the Governance of the Energy Union reiterates the same requirement (Regulation (EU) 2018/1999, Part 3 of Annex V). According to the 2006 IPCC Guidelines, Volume 4, Approach 3 with spatially explicit land-use conversion data (page 3.13) is:

> Approach 3 is characterized by spatially-explicit observations of land-use categories and land-use conversions, often tracking patterns at specific point locations and/or using gridded map products, such as derived from remote sensing imagery. The data may be obtained by various sampling, wall-to-wall mapping techniques, or combination of the two methods. An overview of potential methods for developing Approach 3 datasets is provided in Annex 3A.4.
>
> Approach 3 data can be summarized in tables similar to Tables 3.5 and 3.6. The main advantage of spatially-explicit data is that analysis tools such as Geographic Information Systems can be used to link multiple spatially-explicit data sets (such as those used for stratification) and describe in detail the conditions on a particular piece of land prior to and after a land-use conversion. This analytical capacity can improve emissions estimates by better aligning land-use categories (and conversions) with strata mapped for classification of carbon stocks and emission factors by soil type, vegetation type. This may be particularly applicable for Tier 3 emission estimation methodologies. However, issues of compatible and comparable spatial resolutions need to be taken into account.

In Annex 3A.4 to the Volume 4, further guidance is given. When Approach 3 is used, the country is divided into grid cells or small polygons if they are spatial units. Land use and land-use changes to the spatial units can be estimated by sampling or wall-to-wall methods. The Guidelines points out possible misclassification when land cover data are used for land-use categories. In this study possible misclassification was detected but not corrected. The total land and inland water areas were scaled separately for Southern and Northern Finland to the total land and water areas which were used in the GHGI and which are announced by National Land Survey of Finland. The correction done due to misclassification was to use the data only from forest land sample plots to estimate carbon stock changes, thus, for example, misclassification of croplands to forest land were ignored.

Conclusions

The methods applied in the Finnish GHG inventory have developed over the time; new land categories and emissions and removals related to them have been included to increase the coverage and quality. The sample-based system for area estimation has been functional, and still is. Disadvantages of the current method falls upon lack of data for the recent years, and detection of land use and land-use change whose areas are small. The recalculations in the coming years change the situation, sometimes very radically, which is undesirable. Therefore, it is important to examine the possibilities of new data sources, like Copernicus CLC+ data. Applying the Finnish CORINE Land Cover data to classify lands into the IPCC land use categories brought up some basic problems. Some issues relative to GHGI are listed below.

Different classifications: CLC is land cover data and in the IPCC classification the land use is the determining factor and the land cover secondary. This was shown especially on forest land and settlements. Settlements with forest cover were classified as forest land, producing a low area for settlements.

Misclassification: Grassland areas were at the same level on both datasets, but contents were different as abandoned fields with some trees tended to be classified as forest land in Corine and less intensively cultivated fields were still in Cropland category in GHGI and grasslands in CORINE. Separating peatland forests from other forest land was also challenging according to the analyses.

Emissions and removals: The area data need to be linked to appropriate carbon stocks and emission factors. When area estimates and carbon stock changes are estimated based on different data sources, there is a risk to break this linkage. For example, forest land area is estimated from CLC data and tree biomass stocks from NFI data. The third dimension to the estimation is the soil type and drainage sit-uation on organic soils. *Time series*: CLC data covers years from 2000, meaning that the first ten years of the time series are lacking, but required for the GHGI. Further work is needed to solve this problem, keeping in mind the time series consistency. *Placing events for right years*: Annual land-use change areas are averages between two CLC rounds. The Finnish GHGI aims to report annual changes in occurrence year, especially afforestation and deforestation. *Uncertainty*: All above mentioned issues, in addition to other matters, should be considered when the uncertainties are estimated.

The current Finnish method employs georeferenced data combined with ground survey data (NFI) to produce land use and land-use change areas for regions. This feature is used to report afforestation and deforestation areas and emissions under the Kyoto Protocol, albeit the results are presented for two combined areas, Southern and Northern Finland (Statistics Finland 2020). However, development work is needed to

- establish reliable monitoring system which observes also small or rare occurrences in land use and land-use change.
- establish reliable land use and land-use change estimates for smaller regions than at present.
- utilize current and new data sources to decrease uncertainties.
- build a verification procedure to compare GHGI areas with an independent monitoring system, for example, Copernicus (if not used in GHGI)
- build a verification procedure to compare GHGI emission and removal estimates with in-dependent assessment, for example, atmospheric measurements.

Testing different top-down approaches for CO_2

Additionally, the Finnish partners (FMI and SYKE) tested novel data fusion solutions where they applied national land cover datasets for complementary approaches to monitoring LULUCF related emissions and removals and their temporal evolution. These complementary approaches build on the use of atmospheric greenhouse gas data and include top-down approaches by atmospheric inverse modelling and top-down approaches that directly utilise atmospheric satellite observations (including Copernicus products).

As an example of applying atmospheric inverse modelling data, reanalysis data products of CO_2 fluxes were taken from the Copernicus Atmosphere Monitoring Service (CAMS, version v18r3, in situ assimilation) and analysed over Finland. Attempts to analyse CO_2 fluxes from the different land categories were not particularly successful due to the too coarse spatial resolution of the CAMS reanalysis data for this purpose (3.75 degrees in longitude and 1.875 degrees in latitude). Total fluxes were nevertheless aggregated over Finland and found to differ significantly from the national GHG inventories. Interannual variability in the optimised biospheric fluxes from CAMS was found large over Finland. Additional methods for studying the CO_2 uptake were therefore considered important.

Another complementary approach was developed with the use of atmospheric satellite observations in order to study the possibilities of novel Earth Observations as support to LULUCF reporting. Satellite observations of greenhouse gases describe the total column abundance of CO_2 or CH_4. Therefore, these are always affected by atmospheric transport and cannot be directly interpreted as emissions from the region of observation. However, the satellite missions designed for retrieving these gases also provide observations of the solar-induced chlorophyll fluorescence (SIF). SIF is a reasonably novel satellite product and describes the photosynthetic activity of vegetation. It can be used to estimate gross primary productivity (GPP) on the canopy level and thus link with e.g. land surface models that estimate natural CO_2 fluxes. It was found in the project that satellite SIF data can be used to distinguish and analyse differences in the uptake of different land cover classes and their seasonal and interannual variability. In particular, agricultural lands were found to have the highest annual SIF-GPP estimate. **This method is scalable to other countries and the pan-EU region and can in the near future greatly benefit from the SIF data from Sentinel 5P TROPOMI.**

Methane emissions from Finland

The Finnish national anthropogenic methane emissions are reported by Statistics Finland and include emissions from agriculture, waste, energy and industry sectors, and LULUCF emission category as prepared by LUKE. The reporting follows IPCC 2006 reporting guidelines with 2019 refinement. The LULUCF values prepared by LUKE for this project and other emission categories obtained from Statistics Finland (OSF 2021) are presented in *Figure 3.17*. The total emissions incl. LULUCF indicate slightly decreasing trend for the period of 2005 - 2018. The total emissions varied between 210 and 280 kt(CH_4)/yr during the period. LULUCF emissions also had a decreasing trend, and they varied between 26 and 48 kt(CH_4)/yr depending on the calculation method, CORINE-land cover-based method giving smaller values than GHGI-based method.

Methane emissions can be detected using atmospheric concentration data, and the methodology can be used to estimate national budgets, provided the simulations are made in high resolution and the coverage of atmospheric observations is adequate. The FMI atmospheric inverse model Carbon Tracker Europe – CH_4 (CTE-CH_4, Tsuruta et al., 2017) estimates methane emissions in high 1x1 degree resolution, using measurements of atmospheric methane concentrations to constrain land surface emission estimates. The system consists of prior flux maps for natural and anthropogenic emissions, an atmospheric transport model and EnKF data-assimilation scheme. Here we use observations from a global in situ measurement network (six stations in Finland), and S5P-TROPOMI satellite column data. Currently CTE-CH_4 solves emissions in 1x1 degree (~100km) resolution over the northern Fennoscandia and in weekly time step for the period of years 2000-2018.

The atmospheric inversions use anthropogenic and natural emission priors in their emission calculations. The anthropogenic priors were obtained from TNO (Kuenen et al. 2014, Denier van der Gon, 2017) and EDGAR5 (Crippa et al., 2019) gridded global databases. TNO emissions for Finland (187 – 224 kt(CH_4)/yr) are consistent with those reported by Statistics Finland without the LULUCF category, and have a decreasing trend. In posterior the emissions are slightly increased (210 - 250 kt(CH_4/yr)) but still relatively close to those reported by Statistics Finland.

EDGAR5 reports emissions consistently with the IPCC2006 reporting guidelines. EDGAR5 does not include LULUCF emissions as separate source category. For Finland EDGAR5 shows significantly larger emissions than those estimated by Statistics Finland (without LULUCF). EDGAR5 emissions vary from 760 kt(CH_4)/yr to 920

$kt(CH_4)/yr$ compared to around 200 $kt(CH_4)/yr$ by Statistics Finland, and they have a slightly increasing trend after 2004. In posterior the anthropogenic emissions vary from 860 kt(CH4)/yr to 520 $kt(CH_4)/yr$ approaching levels in earlier EDGAR release, EDGAR4.3.2. The smallest emissions occur during the latest years because the emissions decrease, or possibly because the atmospheric observation network is extended affecting the way how the emissions are constrained. Is seems quite evident, though, that EDGAR5 emission estimates are too large for Finland.

The natural prior emissions are obtained from ecosystem models, here JSBACH-HIMMELI (Raivonen et al., 2017, Kleinen et al., 2020) and LPX-Bern v1.4 (Spahni et al., 2012, Stocker et al., 2014). They include emissions from peatlands and mineral lands. In case of JSBACH, peatlands refer to open bogs and inland marshes in EU-Corine land cover and mineral lands to all other land, including managed and unmanaged forest and non-forest lands. Mineral lands emit methane if the average soil moisture in a grid cell is high enough and act as a sink of methane if the soil is dry. Anoxic organic soil conditions and carbon input from vegetation induce methane emissions in peatlands. The short-term emission variation is driven by soil water table level and temperature. Emissions from peatlands are much higher per unit area than those from mineral lands, but small areal cover of peat lands may result in lower emissions than those from mineral lands. In Finland, according to JSBACH, peatland emissions strongly dominate over mineral land emissions. LPX-Bern simulates the development of peatlands using a topography driven approach. LPX-Bern also simulates the sources and sinks at mineral lands similarly to JSBACH, as well as emissions from inundated lands (negligible in Finland). As more than 80% (of which 1/4 organic and 3/4 mineral) of land area in Finland belongs to LULUCF forest, cropland, grassland and wetland categories, the mineral land area of the ecosystem models significantly overlaps with LULUCF lands.

In Finland, the JSBACH natural emissions are larger than LPX-Bern emissions (*Figure 3.17*). This difference is decreased in posterior since the natural posterior (peat and mineral land) emissions are decreased from JSBACH prior and increased from LPX-Bern prior, ranging from 220 kt(CH4)/yr to 480 kt(CH4)/yr. There is no clear trend in the posterior natural emissions. In comparison to LULUCF emissions both models predict significantly (about 10x) larger emissions, which is expected, as they include methane emissions from natural wet peatlands. Anthropogenic emissions from EDGAR dominate the country total emissions in the prior, but in posterior the share of natural emissions is well comparable to anthropogenic emissions. Especially, TNO emissions are of similar magnitude or even smaller than natural emissions in posterior. However, it is noted that a minor part of the natural emissions probably can be attributed to LULUCF category. If we subtract the natural wet peatland emissions from the natural posterior emissions, we get values ranging roughly from zero to 50 $kt(CH_4)/yr$, which would thus be related to emissions from managed and non-managed forest lands, croplands and grasslands. However, we cannot be sure if the relative proportions of the emissions are the same in the posterior than in the prior.

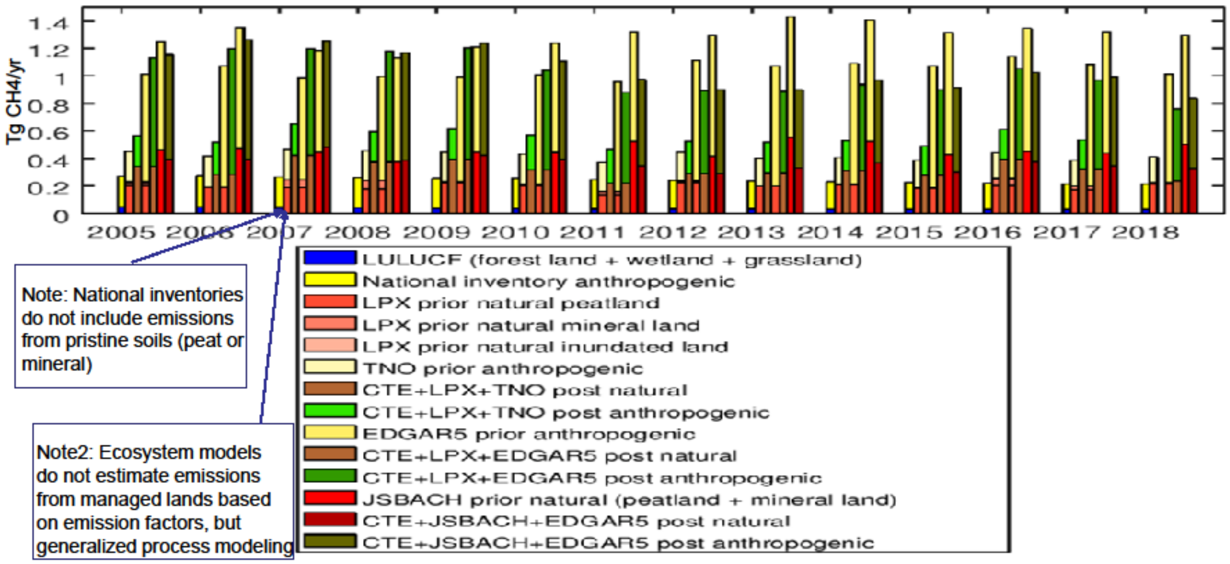
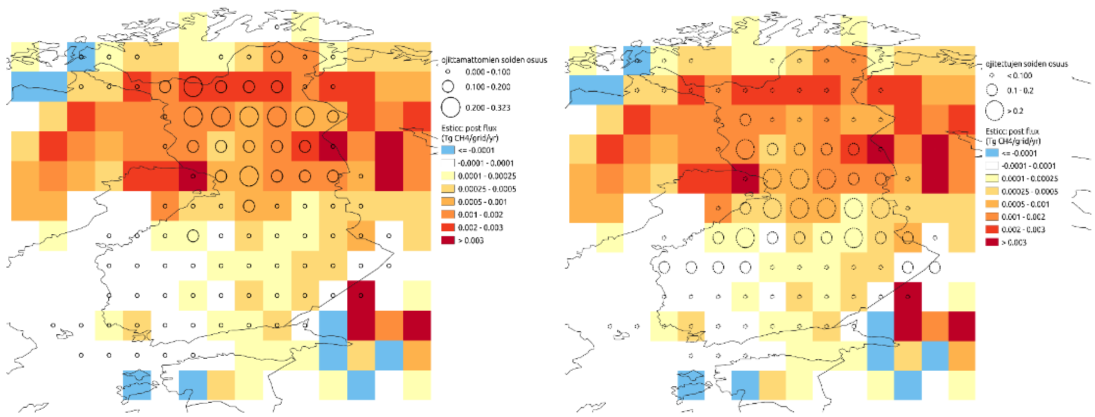


Figure 3.17: Anthropogenic and natural annual methane emissions from Finland, estimated by national inventory and other gridded inventory databases (TNO, EDGAR), ecosystem models (JSBACH, LPX-Bern) and inversions (CTE).

Methane emissions from different land use categories

More than half (58%) of the 8.3 million hectares of peatlands in Finland were drained in the mid 1900's for needs of wood production. Harvesting of peatland forests is continuing even today, though timber production is low on almost one fifth of those forests, mainly due to nutrient poor soil. Harvested peatlands may potentially return to wetland state, if the drainage ditches are not maintained. Drained peatlands emit mainly carbon dioxide by decomposition of the peat, while undrained water-logged peatlands may eventually become carbon sinks but emit methane, which is more potent greenhouse gas than carbon dioxide. However, drained peatlands can also act as a carbon sink if they are rich in growing-phase trees that store carbon. Here we focus on methane emissions from peatlands. Natural peatlands, drained peatlands, forested peatlands, and peat production areas are included in Finnish land cover data as presented by SYKE, and greenhouse gas inventory data is provided for the managed peatland areas by LUKE.

The peatland classes in Finnish national Corine Land Cover were combined with gridded methane emission estimates (*Figure 3.18*). The largest methane emissions occur in the northern Finland, occupied by natural peatlands, while the most extensive peatland drainage areas are in the middle of Finland.



A B

Figure 3.18: A) Natural methane emissions from CTE-CH_4 (color scale) and un-drained peatland fraction (circles) B) Natural methane emissions (color scale) and drained peatland fraction (circles) over Finland. Methane emissions are averaged over years 2010-2012.

We studied the use of Copernicus data to support identification of those land use categories which have large methane emissions. Copernicus Land Water/Wetness product (Copernicus WAW) used here is a compilation of data from 2012-2018 and presents a mapped average wetness distribution over Europe. Here we show results for Finland, **but the approach will be applied in the pan-European context and reported later**. Generally, there was a good correlation with the permanently wet area from Copernicus Land Water/Wetness product and the natural methane emissions from CTE-CH_4 (*Figure 3.19*). It is noteworthy that the correlation improved in the posterior, suggesting that the Copernicus data was able to identify the wet, methane emitting land areas matching with the posterior inversion estimates. The highest emissions occurred in the areas where, as expected, open bog area was largest. High emissions also occurred together with large areas of forest land and transitional woodland, but these may partly be due to open bogs residing in the same grid cell. It is thus challenging to separate the different land use categories from the inversion results, as the resolution is still rather coarse. However, the use of Copernicus WAW data together with methane inversions shows promise for identification of emissions from different land use categories. Copernicus WAWRA archive including annual data will provide more possibilities for studying year-to-year changes.

The effect of land use change on the methane emissions was also studied, but it was difficult to depict an obvious reason for the emission changes from reporting period to another (2006 –> 2012 -> 2018) based on land cover changes. There are several possible reasons for this. Firstly, the changes in land cover are predominantly small (max 16% of the grid cell area), and there may be counteracting land cover changes in the same grid cell. Land cover changes may also take place in a shorter timeframe than the reporting period and one period can include multiple changes in a certain area. The end-uses would need more work, i.e. to recognize whether the land use change was between high/low emitting class or if only small changes in emissions were expected. Also, the signal can be dominated by unaccounted and rapidly changing anthropogenic source mistakenly interpreted as peatland source. Furthermore, the effect of the climate variability may not be fully considered if the climate responses of the prior ecosystem models are not fully adequate. More investigations on the topic are needed, possibly with longer time series.

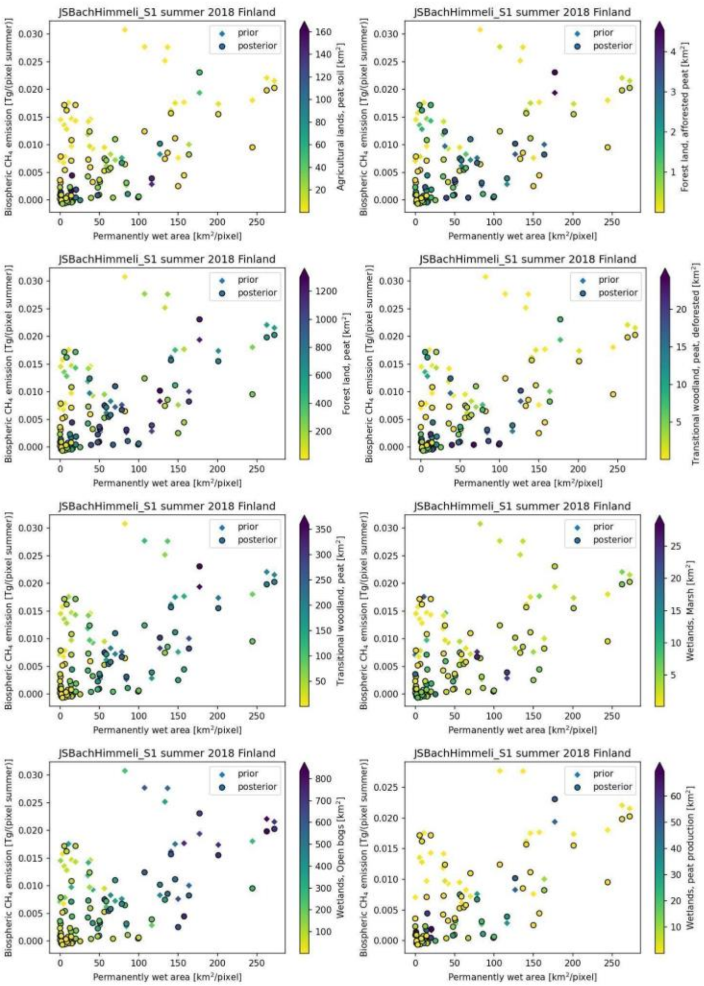


Figure 3.19: Methane emissions from CTE-CH_4 in 1x1 degree resolution over Finland and averaged over June – September 2018, plotted against permanently wet area from Copernicus Land Water/Wetness and area of the specific peat land use (color scale) in the grid cell.

3.3.4. Ireland (MU)

The Irish national inventory agency (Environmental Protection Agency), agreed to work with Maynooth University to identify specific areas that the use of Copernicus Sentinel data may provide additional information or strengthen current national inventory arrangements. Several meetings have taken place and the inventory agency has outlined areas that remote sensing could provide additional support and information that may assist in improving national reporting.

The inventory agency identified the need to validate existing land cover data in order to incorporate Ireland's first complete aerial data survey (PRIME2) into the national inventory data management system. Incorporating a single year's land cover data into the national inventory presents challenges that are common to all inventory agencies. As new technologies become available and data collection capabilities have improved, incorporating this data into the inventory systems will present issues common to all member states (MS) namely; data consistency, the development of appropriate and accurate gap filling methodologies in addition meeting validation requirements as described in the (IPCC,1997).

To develop a solution along with the requirements of the national inventory agency, Maynooth University will apply time series image classification techniques to identify land cover and changes between 2000 and 2018. The aim of this is to provide a methodology that can be implemented by other MS when new data collection activities do not provide a complete temporal coverage. To achieve this, a number of objectives were identified, and include:

- Identification of a test site in which to evaluate the classification process
- Development of machine learning classification workflows to enable other member states to easily customize the process to represent their own land cover
- Provide statements of accuracy using independently validated land cover data
- Provide assessment on the accuracy of CORINE and other CLM data (e.g. HRL).

Table 3.5: Land cover categories according to CORINE land cover dataset.

<u>**CLC Category**</u>	<u>**CLC Sub-Category**</u>
1. Artificial Surfaces	*111 Continuous urban fabric* *112 Discontinuous urban fabric* *121 Industrial or commercial units* *123 Port areas* *131 Mineral extraction sites* *132 Dump sites* *133 Construction Sites* *141 Green urban areas* *142 Sport and leisure facilities*
2. Agricultural areas	*2.1.1 Non-irrigated arable land* *2.3.1 Pastures* *2.4.2 Complex cultivation patterns* *2.4.3 Land principally occupied by agriculture, with significant areas of natural vegetation*
3. Forest and seminatural areas	*3.1.1 Broad-leaved forest* *3.1.2 Coniferous forest*

	3.1.3 Mixed forest *3.2.1 Natural grassland* *3.2.2 Moors and heathland* *3.2.4 Transitional woodland/shrub* *3.3.1 Beaches, dunes, sands* *3.3.3 Sparsely vegetated areas* *3.3.4 Burnt areas*
4. Wetlands	*4.1.1 Inland marshes* *4.1.2 Peatbogs* *4.2.1 Salt marshes*
5. Water bodies	*5.1.2 Water bodies* *5.2.1 Coastal lagoons*

A pilot site was identified by the inventory agency located in Co. Wicklow, Ireland (*Figure 3.20*). The county has a mixture of land cover categories (according to CORINE land cover dataset) described in *Table 3.5*.

This large number of land cover categories makes this pilot site an appropriate location for the inventory agency and Maynooth University to test gap filling and time series consistency and validation procedures. The inventory agency intends to use bottom up data (e.g. Land Parcel Information System, National Forestry Inventory, CORINE datasets) to evaluate agreement between these datasets the PRIME2 dataset.

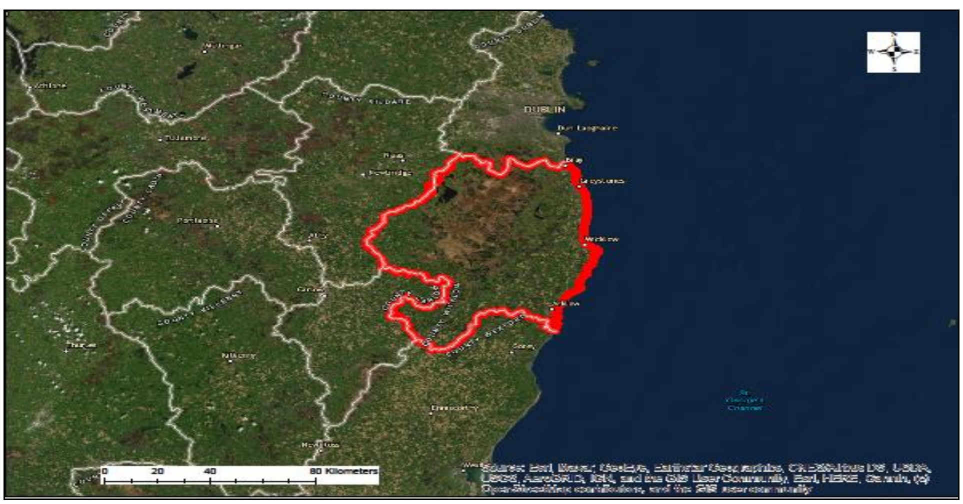


Figure 3.20: EPA pilot site located in the eastern province (Leinster)

Maynooth University are developing methods to supplement the efforts of the agency to understand areas of disagreement between these datasets, to provide higher temporal resolution and to provide ground truth data that has been validated with estimates of accuracy. To achieve this, methods are currently being developed to enable the use of machine learning methods that employ multiple feature extraction techniques and image classification techniques.

Image classification as it relates to land cover estimate requires training data to be provided to train image classifiers. A degree of subjectivity is introduced into this process; however, this was reduced by selected by (1) visually confirming land cover types, (2) validating them against CORINE data and (3) selecting training points where only clear examples of the landcover classes existed (e.g., points that had multiple land cover classes clustered together were not selected).

Additionally, the inclusion of the number of training points was investigated in order to ascertain if it was more appropriate to include highly resolved land use classes (e.g., Continuous or Discontinuous Urban areas as found in CORINE) or more coarse land cover categories (e.g., Settlements as found in the LULUCF emission source categories)

Image classifiers such as Random Forest or Support Vector Machine have been observed to perform better for some land use classes over others. Additionally, using different feature selection methods (e.g., Pixel and object-based methods) have been found to perform better for some landcover classes over others. For instance, Gislason et al. (2006) found that the Random Forest Classifier outperformed the CART classifier when classifying land cover while other studies found Support Vector Machine (SVM) achieving accuracy levels of 90%.

By using several image classification techniques (as shown in *Figure 3.21*) it is hypothesized that the errors related to systematic or random differences can be reduced. The output from this objective may provide important information regarding image classification performance, the best performing combination of feature selectors and classifiers and provide a method to average these results to produce a more accurate classification process. Outputs of this nature are important to determine how to best address the incorporation of new data sources using remote sensing.

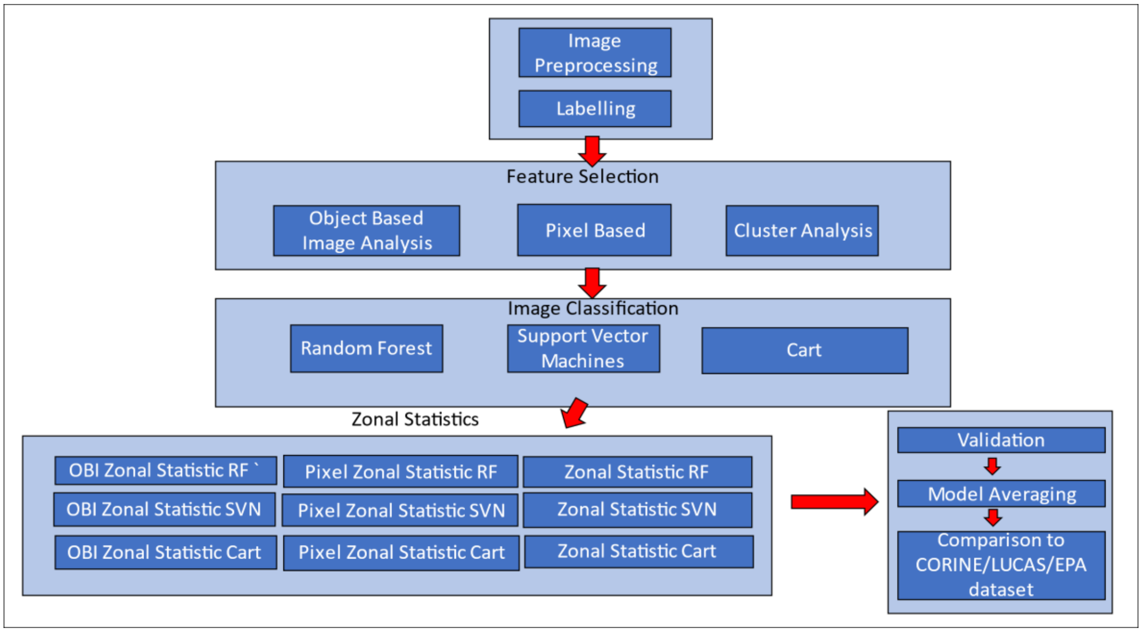


Figure 3.21: Image Classification Process.

The models will then be applied to specific land cover classes identified by the inventory agency and timeseries analysis and zonal statistics will be provided. It is a requirement of the inventory agency that these statistics and methods are documented in a manner which the agency can use during inventory review processes and needs to be easily incorporated into the national inventory system.

These methods will be validated using point source information available from the **Land Use and Coverage Area frame Survey (**LUCAS). This data is available from 2006 at three years reporting intervals.

Preliminary Results

Zonal statistics for landcover classes found under CORINE have been estimated for the following years; 2000, 2006, 2012 and 2018. This data coincides with published versions of CORINE and facilitate a comparison between the two datasets. However, CORINE is likely to have high levels of uncertainty and as a result LUCAS data will be used to validate our image classification analysis. *Figure 3.22* and *3.23* show zonal statistics from 2000 using Landsat 5 data and using two image classifiers (Random Forest and Cart) and classified using an RGB image (denoted Pixel) and using NDVI values. Good agreement was found in 2000 between NDVI and Pixel classifications with the exception of a few land cover classes where large difference were observed (Land principally Occupied by Agriculture).

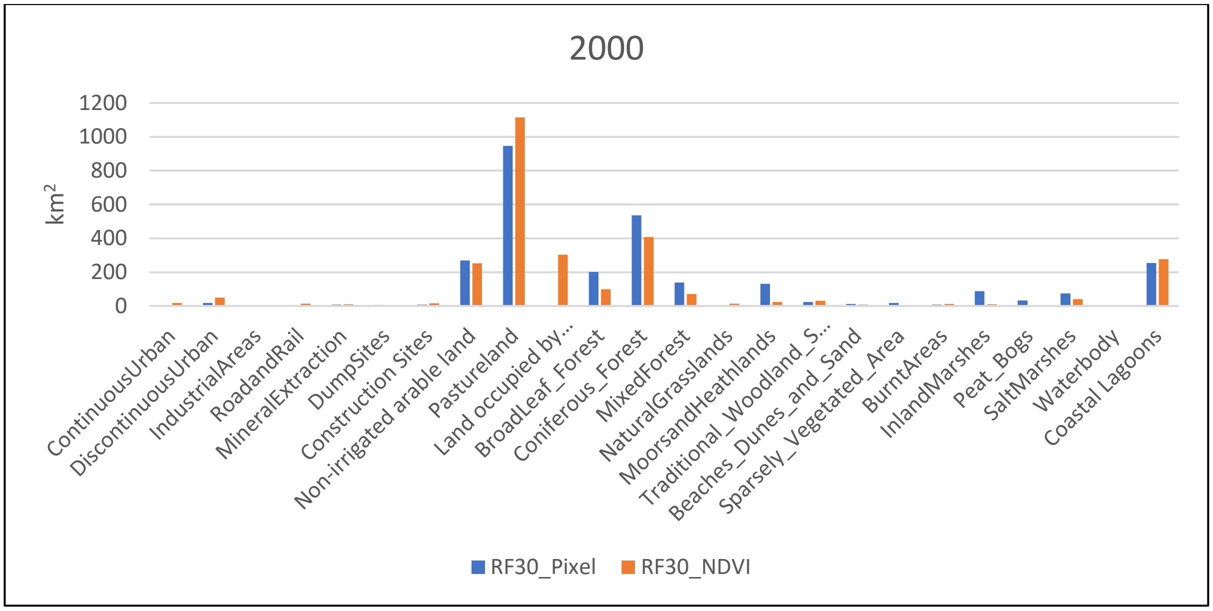


Figure 3.22: Images were classified using Random Forest Image classifier using NDVI and RGB to ascertain level of agreement.

However using the CART classifier was found to have some disagreement between pixel classification and NDVI classification. This indicates that the Random Forest classifier classifies more consistently.

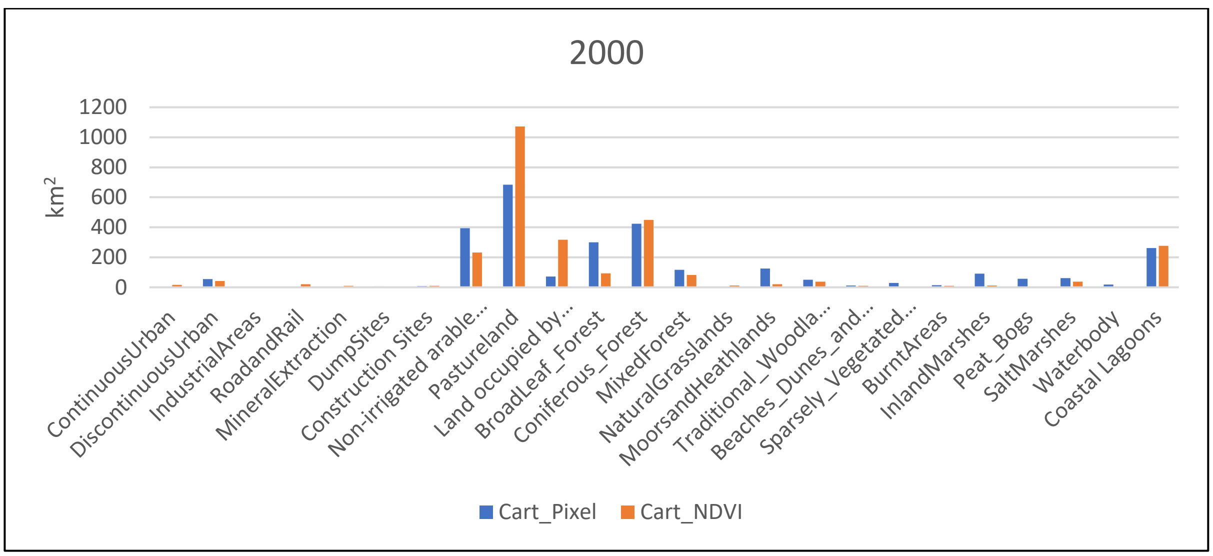


Figure 3.23: Images were classified using Cart Image classifier using NDVI and RGB to ascertain level of agreement.

Finally, classification statistics using different classifiers and using both pixel and NDVI were averaged and compared to CORINE datasets (*Figure 3.24*). Good agreement was observed for important land cover classes like pastureland, non-irrigated arable land, land occupied by agriculture however it poor agreement was observed between the forest categories (Mixed, Broadleaf, Coniferous forests) and Inland Marshes.

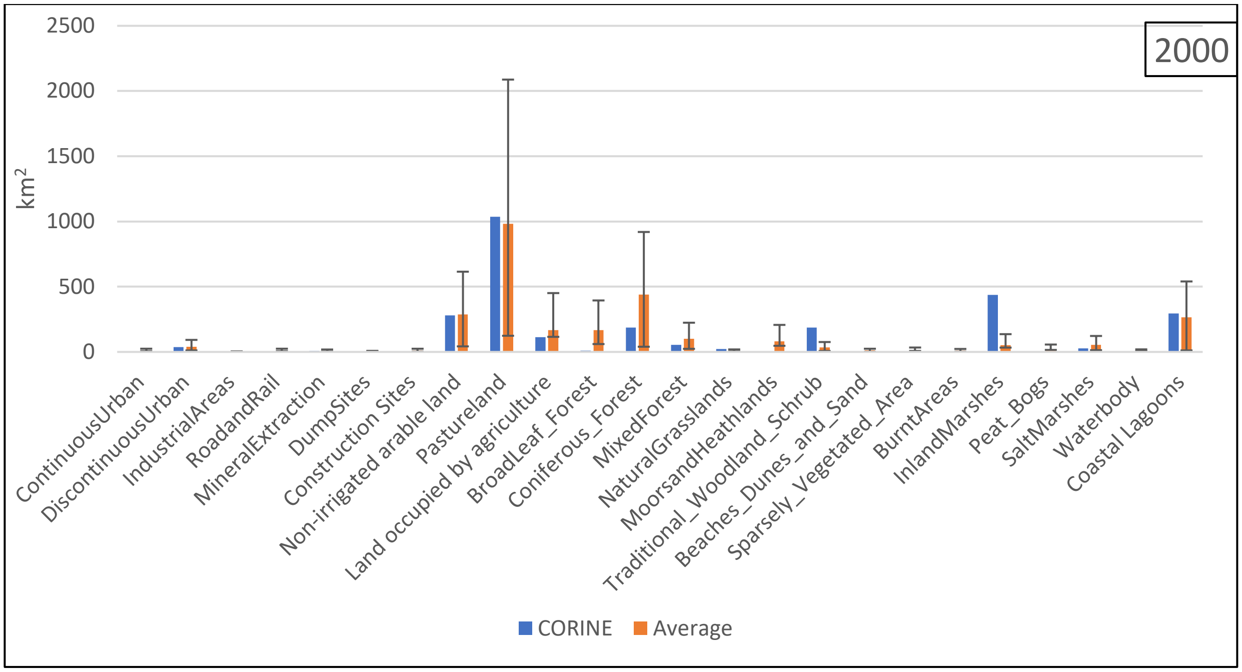


Figure 3.24: Image classifers were averaged and compared to CORINE to ascertain level of agreement.

Lessons Learned

In order for this analysis to be useful to the inventory agency, a statement of accuracy from independent and accurate data source is required. The LUCAS data set is an appropriate data source to achieve this and has good temporal coverage. This will be used to provide an assessment as to the accuracy of the classification techniques.

A full time series needs to be generated using Sentinel and Landsat data and a grid system similar to the methods used by the inventory agency needs to be employed. A common issue associated with using pixel-based methods is the presence of 'salt and pepper noise' however several solutions exist and these will be explored and evaluated in terms of the improvements they deliver to the classification performance.

Further engagement with the inventory agency is then required to better understand how this will be incorporated into the inventory management system.

3.3.5. Poland (IGIK & CBK-PAN)

The study conducted by the team at the Institute of Geodesy and Cartography (IGIK) was preceded by the detailed analysis of the LULUCF requirements, data used, methods and national land use nomenclatures. In addition, a detailed analysis of the currently available Copernicus High Resolution Layers (HRL) was analysed in terms of consistency, accuracy, thematic content, nomenclature, minimum mapping units. The comparison was performed in respect to the LULUCF regulation.

The HRLs for the reference years 2012, 2015 and 2018 were considered and analysed. The IPCC land use categories can be addressed by the following HRLs:

- Forest land = HRL Forest Density Layer
- Grassland = HRL Grassland
- Wetland = HRL Water and Wetness
- Settlements = HRL Imperviousness
- Cropland = Remaining land (none of the above)

However, the comparison of HRL is not straightforward. In the first place, the analysis and adjustment of the HRLs had to be applied to minimise the inconsistency in the datasets and to fulfil the inventory regulations and national definitions. Here are two examples of essential adjustments of the HRLs:

Forest Land is the most important category recognised as the main source of CO_2 removals. The FL category follows the national definition of forest, which refers to woody vegetation with minimum area of 0.1 hectares, minimum width of 10 m and minimum tree crown cover of 10% with trees having a potential to reach a minimum height of 2 m at maturity stage. To meet the criteria of the national definition, the Tree Cover Density Layer could be applied, however, the clear-cuts have to be masked (i.e. included as part of FL) using either the national forest maps or by additional spatio-temporal analysis of a time series of Tree Cover Density Layers. Furthermore, the Forest Additional Support Layer must be applied to exclude the tree in urban context and trees used for agricultural practices.

Another example is the **Grassland** category, where the HRL Grassland can be applied (note, the GRA layer is not available for the reference year 2012). According to the national definition, grasslands include permanent meadows and pastures as well as woody and bushy land. To fulfil the first part of the definition, the grass vegetation including parks, urban green spaces in residential and industrial areas must be excluded from the Grassland Layer. In addition, the Small Woody Features, which is currently available for the year 2015 should be included in the grassland areas.

In the second part of this study, the adjusted HRLs were integrated to obtain the seamless land cover maps. The scheme of the proposed integration of adjusted HRLs into land cover map is presented in the *Figure 3.25*.

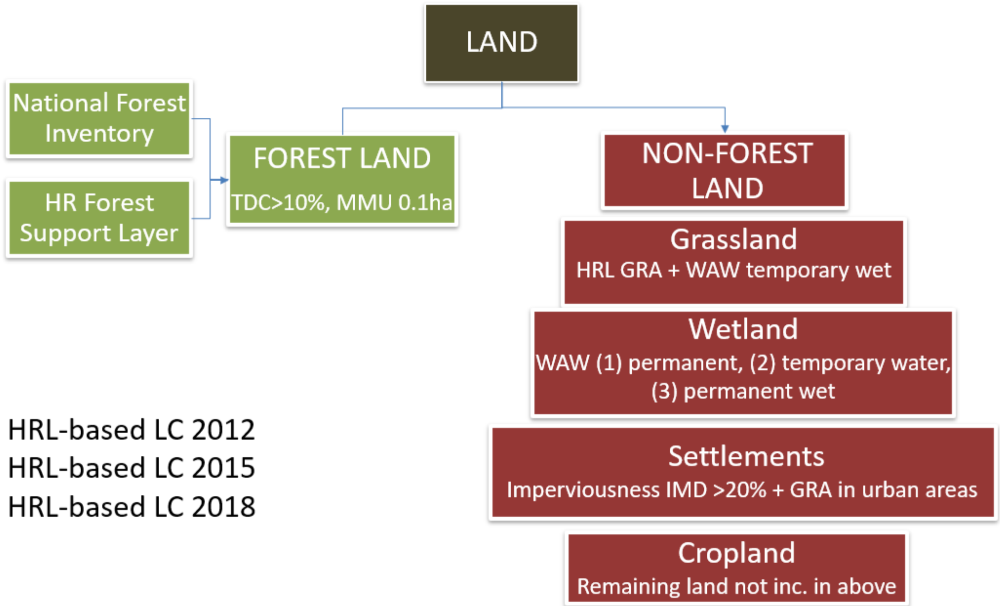


Figure 3.25: Scheme of the integration of adjusted HRL into land cover map.

Based on the developed methodology for adjustment and scheme of integration of the various HRLs, the seamless land cover maps were produced for 2012, 2015 and 2018 for the study area in Poland - NUTS2. *Figure 3.26* presents the land cover map derived based on integration of HRL 2015 for the study area - Podlaskie voivodeship.

Figure 3.26: Example of the land cover map prepared based on the integration of Copernicus HRLs for the Podlaski voivodeship.

The statistics on the area of each land cover type derived from the integration of HRLs for the selected NUTS were calculated. In parallel, the National Centre for Emissions Management (KOBiZE) calculated and provided the area for land cover classes for the selected NUTS for the period 1999 - 2020. In the next step, the HRL's derived land cover areas were compared with the KOBIZE land cover. The result of the comparisons is presented below in *Figure 3.27* (reference year 2012), *Figure 3.28* (reference year 2015), and *Figure 3.29* (reference year 2018).

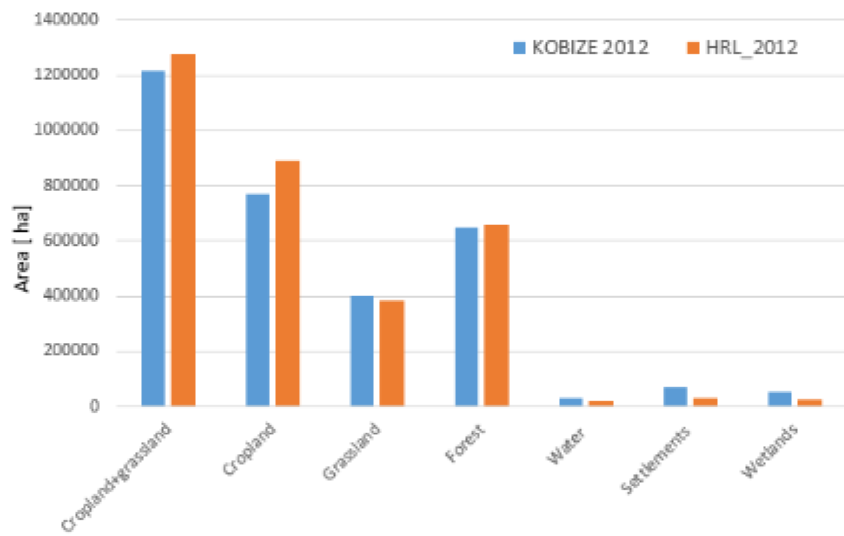


Figure 3.27: Comparison of the total area of land cover classes calculated by KOBIZE and derived based on Copernicus HRLs for 2012 in Podlaskie voivodeship.

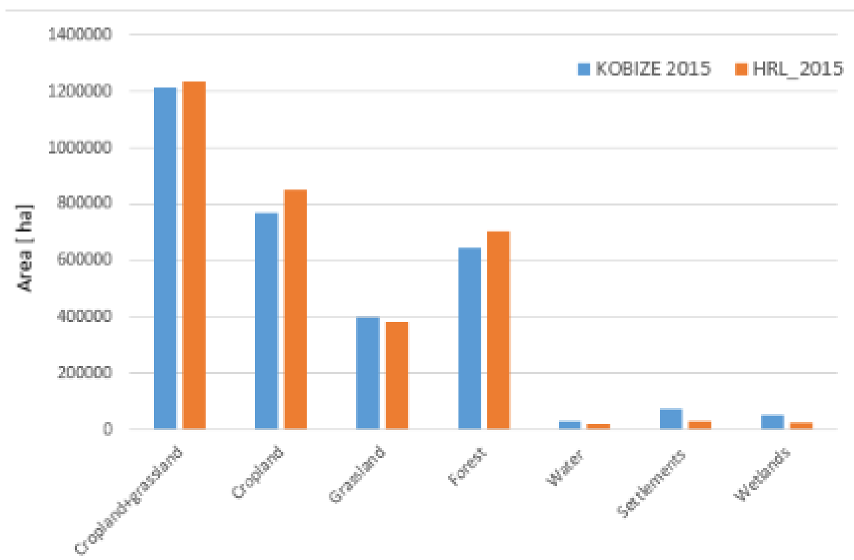


Figure 3.28: Comparison of the total area of land cover classes calculated by KOBIZE and derived from the Copernicus HRLs for 2015 in Podlaskie voivodeship.

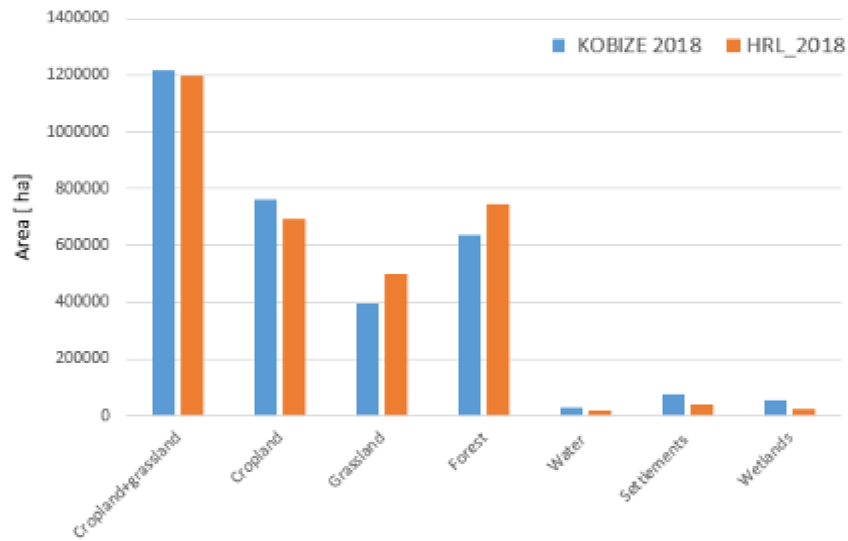


Figure 3.29: Comparison of the total area of land cover classes calculated by KOBIZE and derived from the Copernicus HRLs for the 2018 in Podlaskie voivodeship.

The results show quite good agreement between the land cover areas calculated by KOBIZE and based on Copernicus HRLs. The highest disagreement was observed in case of croplands, where in 2012 and 2015, the HRL showed larger areas compared to 2018, However, if we take into account the total area of cropland and grassland, the areas are comparable in both datasets in three analysed years. The discrepancy in the forest area in 2018, has to be investigated. The area of settlements and wetlands seem to be higher in the KOBIZE dataset.

Table 3.6 presents the proportion of land cover classes in the total area of NUTS2 = 2 018 702 ha.

Table 3.6: Proportion of land cover classes in the total area of NUTS2

	KOBIZE (%)			HRL(%)		
	2012	2015	2018	2012	2015	2018
Cropland+grassland	60,3	60,2	60,4	63,2	61,3	59,2
Cropland	38,2	38,2	37,8	44,2	42,2	34,4
Grassland	19,7	19,8	19,6	19,1	19,1	24,8
Forest	32,1	31,9	31,6	32,7	34,9	36,9
Water	1,4	1,4	1,4	0,9	0,9	1,0
Settlements	3,4	3,5	3,6	1,6	1,6	2,0
Wetlands	2,7	2,7	2,7	1,3	1,2	1,0

Furthermore, the spatio-temporal analysis of the three land cover maps was performed to derive the extent and statistics on the land cover changes between years 2012-2015-2018. *Figure 3.30* presents the results of the land cover change analysis performed based on the integrated HRL land cover.

	2012-2015	2015-2018
A. Forest land		
1. Forest land remaining forest land	629309	679484
2. Land converted to forest land	75718	64682
B. Cropland		
1. Cropland remaining cropland	825790	610687
2. Land converted to cropland	27007	83000
C. Grassland		
1. Grassland remaining grassland	378788	304545
2. Land converted to grassland	6242	195795
D. Wetlands		
1. Wetlands remaining wetlands	40880	24351
2. Land converted to wetlands	2462	16422
E. Settlements		
1. Settlements remaining settlement	32086	27684
2. Land converted to settlements	580	12164

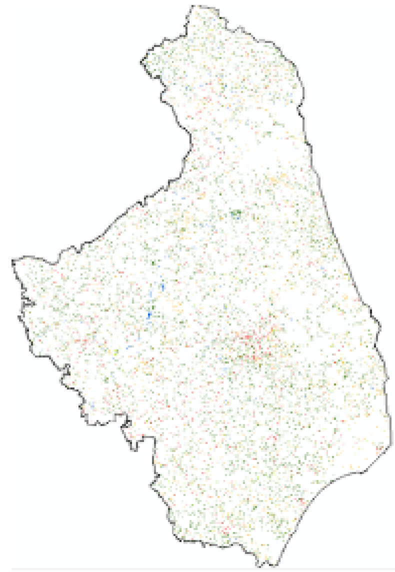


Figure 3.30: Land cover changes between 2012-2015 and 2015-2018 (in hectares) derived based on the land cover maps derived from the integration of Copernicus HRLs.

The largest discrepancy in land cover changes between both periods were observed in areas assigned to land converted to grasslands, croplands, wetlands and settlements, where the change area between 2012-2015 was much smaller than between 2015-2018. The results of this study revealed that many land cover changes derived based on the integration of HRLs are not real changes. This is related to the inconsistency and quality of the HRL, in particular a low quality of the HRL2012. To assess the reliability of the land cover changes derived based on integration of HRLs, the verification of the changes was performed using the national archive aerial orthophotos.

To sum up, the Copernicus Land Monitoring Service products in particular High-Resolution Layers (HRL) are promising for wall-to-wall inventory of LULUCF. However, the following data pre-processing and quality issues must be addressed a) the adjustment of HRLs to the national requirements and definitions of land cover and land use is essential; b) method for Integration of information from the HRLs should be preceded by deep analysis of GHG emissions and removals methodology; c) quality and accuracy of HRLs must be considered and further improvement of the systematic omission and commission errors in HRLs have to be addressed. It must remember that the methodology for LULUCF must be reliable, accurate and repeatable.

Land Cover classification of the Podlaski voivodeship performed using S2GLC approach (CBK PAN)

S2GLC approach has been used for LC classification of Podlaski voivodeship. The territory of the voivodeship is covered by six Sentinel-2 image tiles. According to the classification procedure for each tile, a set of multitemporal images was collected and then classified using Corine and HRL as references. The final results were mosaicking into one file and also intersect with voivodeships borders. The classification was done for the year 2017, 2018, 2019 and 2020 (*Figure 3.31*). Due to the availability of Sentinel-2 data, it was impossible to process images from previous years. For the same reason, the S2GLC classification does not match the dates of HRL's production. Only the year 2018 can be considered for comparison, comparison of change detection cannot be performed like it was demonstrated based on HRL data.

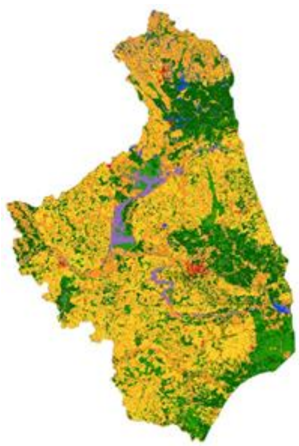
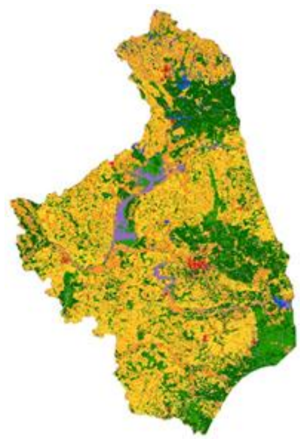
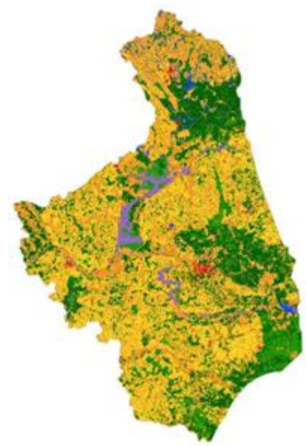
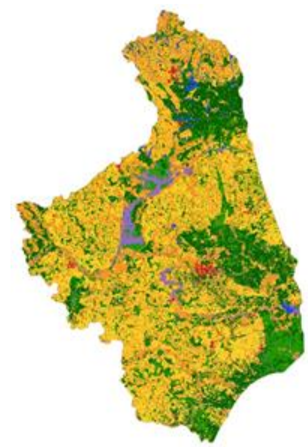
LC 2017	LC 2018	LC 2019	LC 2020
			

Figure 3.31: S2GLC Land cover classification of Podlaski voivodeship.

Results with division into 9 land cover classes are presented in *Table 3.7*. Based on it two graphs were prepared to show trends of changes of "Artificial surfaces and constructions" (*Figure 3.32*) and "Forest" class which is a sum of "Deciduous broadleaf tree cover" and "Evergreen coniferous tree cover" (*Figure 3.33*). In the first case, the area covered a built-up area slightly but steadily increasing. On the other hand, the area of "Forest" in the years 2017, 2018, 2019 decreased and then there is an upward trend in 2020. Both graphs show a situation that seems very likely.

When analysing the results of the S2GLC classification, it should be remembered that image processing is completely automatic and the approach is designed for production LC once per year. The cost of S2GLC is incomparably lower compared to Corine CL or HRL which are produced respectively once per 6 and 3 years. The area of the whole EU Europe can be processed during 6-7 weeks, and one week is enough for classification a country of size like Poland. Additionally, the final resolution of land cover is 10m. S2GLC is a good complement to CORINE LC and HRL databases.

S2GLC approach is systematically improved. Recently, CBK PAN works on the new change detection method, which is adapted to the classification algorithm.

Table 3.7: Results of LC classification performed for year 2017, 2018,2019, 2020 using S2GLC approach.

LC class	ha			
	2017	2018	2019	2020
Excluded (clouds)	0	0	0	1
Artificial surfaces and constructions	24936	25803	27365	28737
Cultivated and managed areas	598075	619059	621281	607558
Deciduous broadleaf tree cover	297931	292850	288548	286582
Evergreen coniferous tree cover	440385	436053	439600	444230
Herbaceous vegetation	608500	587238	588651	604464
Marshes	29419	37816	33213	26826
Peatbogs	927	797	755	1685
Un-Consolidated areas	569	602	450	722
Water bodies	17203	17727	18081	17141
sum	2017945	2017945	2017945	2017945

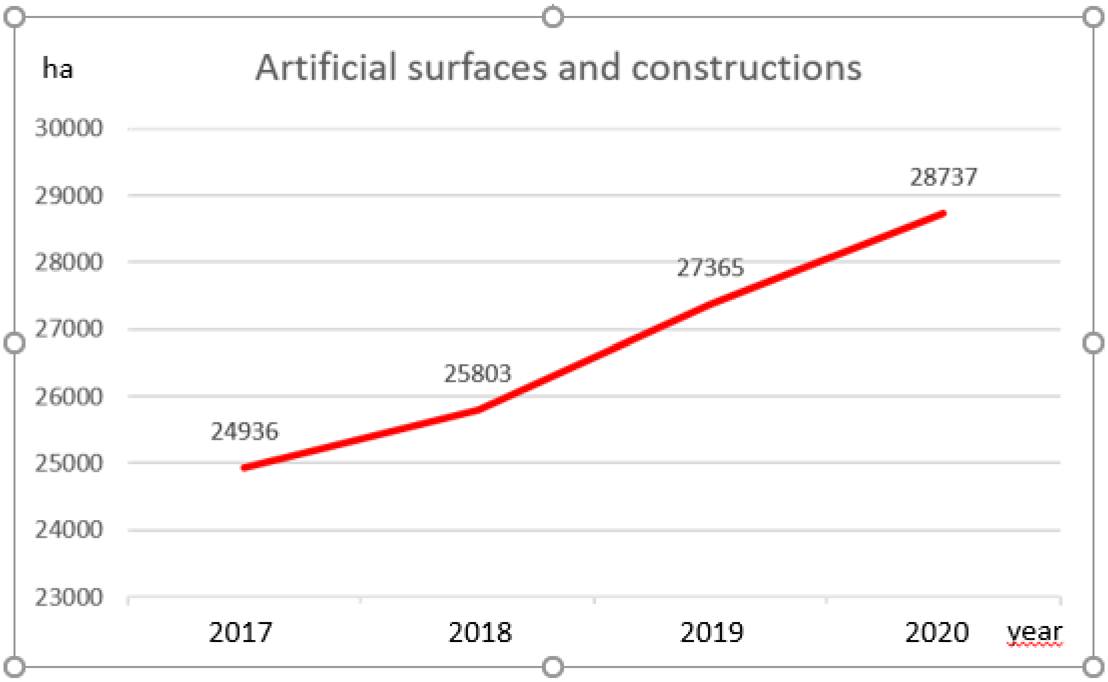


Figure 3.32: Trend of build-up areas based on S2GLC classification - Podlaski voivodeship.

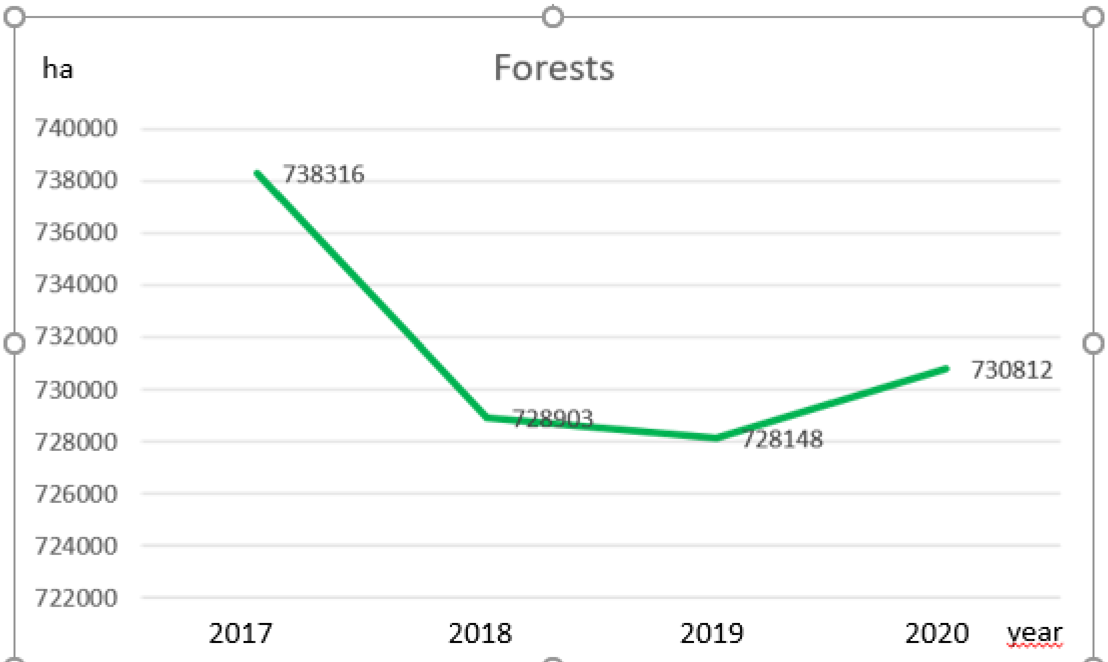


Figure 3.33: Trend of forests based on S2GLC classification - Podlaski voivodeship.

3.3.6. Spain (IHCANTABRIA)

Through the search for Copernicus solutions to the gaps detected in Spain, methods and approaches to monitor and report gap-related information have been identified. These proposed solutions will allow advancing the IPCC Approach and Tier currently achieved by the Spanish Inventory.

The currently available cartographic system of Spain aims following an IPCC Approach 3 (geospatially-explicit). At this regard, the Spanish LULUCF inventory has recently made an effort to develop a geographically explicit land use change monitoring system according to IPCC Approach 3 (geospatially-explicit) from 1970 to the present. This system has been based on the collection and homogenization of all the existing historical cartography of land use and land cover (forestry, crops, CORINE, SIGPAC, etc) from 1970 to 2015. The results of this new project are going to be used in the calculation of the emissions and removals of the LULUCF sector in the next 2020 edition of the National Emissions Inventory. One step beyond, the National Inventory System is currently evaluating/exploring different alternatives for including the temporal dimension into the actual geographically-explicit land use-land cover change monitoring system, for the foreseen 2015-2030 monitoring period. Among the possible options, this project has fostered assessing the capabilities, tools and datasets from the Copernicus program.

The combination of Regional or Member State-specific data needs for monitoring emissions and removals from land use, land use change and forestry, together with the existing and forthcoming Copernicus data and services will be an opportunity to evaluate the data coherence and compatibility that may exist at a national level between historical cartography and Copernicus products. Added to the evaluation of such degree of compatibility, another result of the project has been the identification of the main obstacles and gaps, and so will be considered the technical solutions necessary for a robust and coherent articulation between the two systems.

In addition, the Spanish inventory is seeking a Tier 2 (T2) in the categories cropland (CL), grassland (GL) and wetland (WL) in the Mediterranean domain, providing also consistency with SIGPAC database. The Spanish LULUCF Inventory has a recognized experience in the application of Tier 2 methodological levels with "country specific" data for certain use transitions or even remaining uses (i.e. Cropland remaining cropland for the case of transitions involving a woody crop and particularly olive groves and vineyard). However, for the uses of CL, GL and WL that represent an area of 323,000 km2 (approximately 64% of the national territory), Tier 1 methodologies are often applied for the estimation of emissions and absorptions (E/A) linked to changes in live biomass (LB), and organic carbon in the soil (SOC) in the remaining uses, assuming a neutral balance. In this sense, Copernicus tools are considered to be able provide very useful information regarding management practices of these soils, or the temporal evolution of their natural state that could allow better inferred carbon gains or losses by applying Tier 2 methodologies.

The project analyse different technical options to improve the E/A estimations using T2 methodologies in uses such as CL, GL, and WL. In addition, such analysis at the Spanish national level would imply the evaluation of different climatic areas (Mediterranean, humid temperate or insular), with their intrinsic characteristics. The technological solutions that are proposed as a result of the project can be extrapolated to other countries in the Atlantic and Mediterranean areas. The results of the 6 success stories are presented below. These success case studies in selected NUTS correspond to the most representative solutions that have been deployed for all gaps identified during the participatory processes. They will be also the case studies that will be presented during June and July in the final workshops at both the MS and EU levels.

Success story #1 for LULUCF report in Spain. Wetland monitoring success story: changes in LULUCF and surface water dynamics

Keywords: Modified Normalized Difference Water Index/ Wetlands/ Surface water dynamics/ LULUCF/ SAR data/ Optical data/ Sentinel 1/ Landsat

Application field: changes in surface water dynamics for LULUCF reporting.

Introduction

The anoxic conditions that can be found in wetlands make wetlands an optimal natural environment for carbon sequestration from the atmosphere (Mitsch *et al.*, 2013). An example of the important role of wetlands as atmospheric carbon sinks is that even though wetlands cover only 4-6% of the Earth's land area (Grasset *et al.*, 2017; Strachan *et al.*, 2015), their carbon sink capacity accounts for 20% of the total carbon in terrestrial ecosystems (Carnell *et al.*, 2018; Mitsch *et al.*, 2013), which is much higher than would be expected given the area they cover.

In wetland ecosystems, the complexity of the system, its water conditions and dynamics can vary tremendously, in terms of the timing and duration of the superficial water filling, as well as seasonal patterns of inundation or the transitional land between terrestrial and aquatic systems (Warwick & Brock, 2003). This affects the carbon sequestration capacity of wetlands and should therefore be reported.

From the remote sensing side, this work identifies some tools that could support management and monitoring of the wetland, tests them and proposes them as reliable tools for use for management purposes. In this case, the wetland to be studied is Fuente de Piedra, in Malaga.

The first objective of this study is to tests the level of reliability of applying EO techniques and information management tools like GIS and Remote Sensing (RS) in supporting monitoring short term (seasonality) and long terms (annual) LULC change in Fuente de la Piedra case study. Different types of satellite images are assessed, including the latest products of Sentinel images of the Copernicus program of the European Commission and Landsat images, comparing the validity and effectiveness of the two information sources.

The other objective is to assess of the Surface Water Dynamic (SWD) product in Fuente de Piedra case study, in order to assess its applicability providing accurate open water surface estimations and take a step further in enhancing water dynamics monitoring in wetland ecosystems. The developed SWD methodology assesses the extension of open water areas in wetland ecosystems, which can support and connect with further provisioning services. The product considered is the SWD Temporal Frequency (TF), which describes the number of submersion occurrences relative to the number of image acquisitions during the study period in percentage. Two layers were generated for Fuente de Piedra wetland: one using optical satellite data (Landsat time series), covering the period between January 2007 and September 2015, and the other using Synthetic Aperture Radar (SAR) data (Sentinel 1 time series), from November 2014 to March 2017.

Study area

Fuente de Piedra is a Mediterranean seasonal saline lake located in the northwest of the province of Malaga, region of Andalucia, in south Spain, covering an area of almost 1400 ha. The wetland has shallow, salt water between autumn and spring, and is internationally recognized as home to the largest colony of flamingos on the Iberian Peninsula and the second largest in Europe (de Andalucia, 2012). Its importance also lies in the fact that it has been declared a Ramsar site in 1983 and as nature reserve from 1989, being recognized as Natura 2000 site, Special Protection Area (SPA) and Site of Community Importance (SCI).

The lake is close to the wetland areas of Campillos (formed by six small lakes) and the lake of La Ratosa, which together constitute a wetland complex with similar characteristics in terms of its origin and animal and plant communities.

As shown in Fig. 3.34, the delimitation of the wetland covers the waterbody as set by Ramsar while the Natura 2000 delimitation also covers a more extensive area around the lake. However, this delimitation does not consider strictly the ecological nor hydrological settings, but it is rather influenced by administrative boundaries, related to infrastructure (urban cores, roads, railways) and economic activities (industrial and agricultural areas).

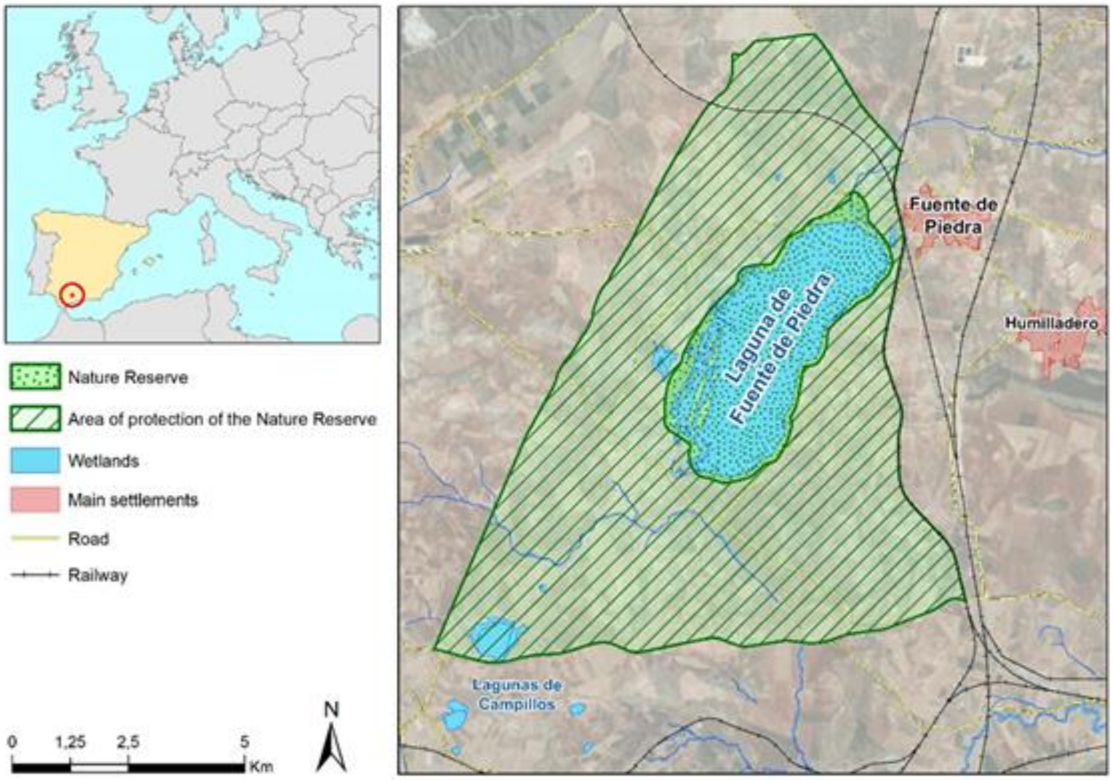


Figure 3.34: Location and delimitation of Fuente de Piedra wetland and its protected area.

The wetland owes its water levels to precipitation runoff and underground water table of the endorheic hydrological basin in which it is located (Rodriguez-Rodriguez *et al*., 2016). The croplands around the wetland consume large amounts of water resources that affect the quantity and quality of the water underneath (González Báez *et al*., 2017; Martos-Rosillo *et al*., 2017). The changes in land uses in the last decades and the expansion of agricultural practices in and around the boundaries are an important pressure on water resources in the area and a major driver of the change in the level of groundwater (Catalán *et al*., 2002). Due to these circumstances, it is necessary to apply a wider delimitation of the study area than either the water body or the administrative delimitation. Such delimitation approach need to ensure that the study area covers the hydro-ecological setting of the wetland including the vegetative structure, the aquatic barriers, wetness, flow gradient, flow volume, flow regime, and protection and political instruments in place supporting environmental conservation, including the Ramsar Convention, the Natura 2000 site, and surface and groundwater water policies.

Figure 3.35 shows the limits of the area of hydrologic processes (hydrological and hydrogeological basins) of Fuente de Piedra and the direction of surface and groundwater flows. These limits are set by the water authorities of Andalucia and are based on hydrological and hydrogeological studies. It is clearly visible that these limits do not correspond to the delimitation of the protected area set. Therefore, the delimitations of both the protection and hydrologic processes were combined in order to consider the entire ecosystem functional area and produce more useful LULC maps for the wetland managers (Abdul Malak *et al*., 2016).

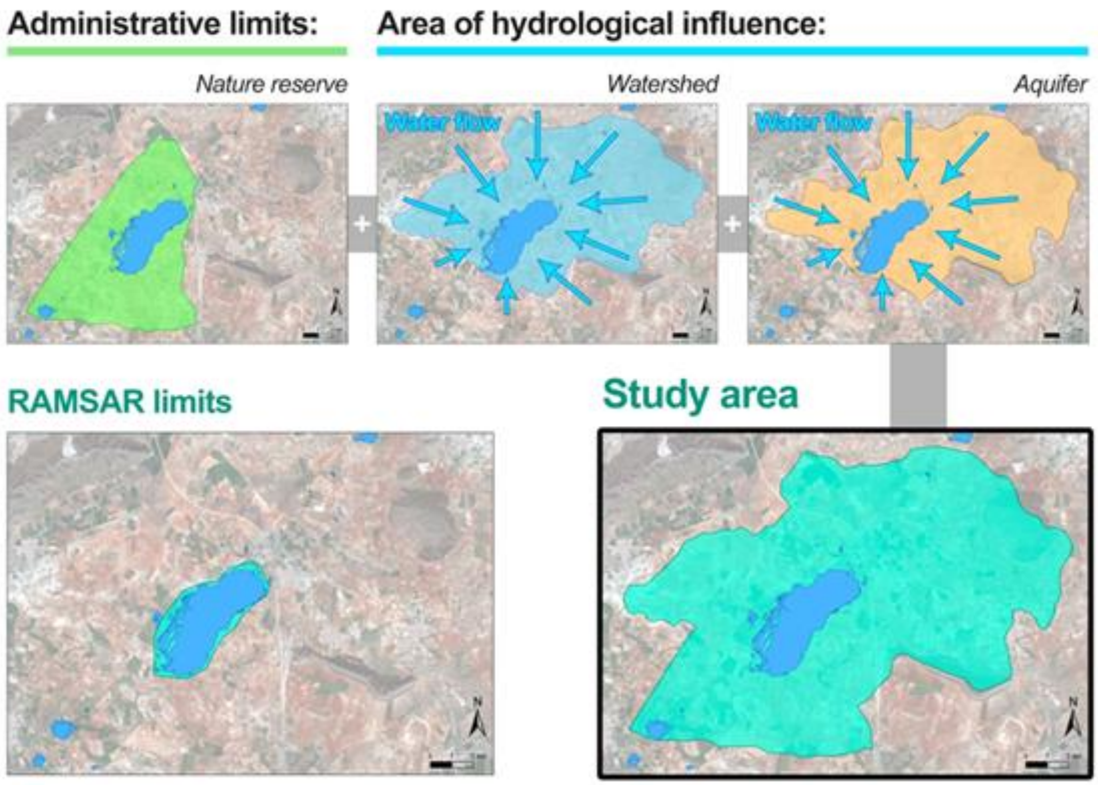


Figure 3.35: Delimitation of the study area according to hydro-ecological setting and administrative limits of the nature reserve.

Monitoring wetlands LULUCF changes

Material and methods

Landsat 8 and Sentinel-2 imagery from autumn 2015 to summer 2016 were used for the LULC classifications to cover the period of maximum flooding in the wetland. It was attempted to collect images on similar dates so that the images of both satellites showed the same information on the land surface.

Images where calibrated and atmospherically corrected to get bottom of atmosphere reflectance. In addition to the corresponding calibration and atmospheric correction, the 20 m bands of the S2 imagery were resampled to 10 m in order to harmonize all band data and ensure that the segmentation of S2 images was adjusted to this resolution.

LULC maps of Andalucia region were used as reference data to assess the accuracy of LULC classifications. These maps were produced by the regional government, Junta de Andalucia, and are mainly based on photointerpretation of high-resolution satellite imagery and aerial photographs with some field work validation. The maps have a scale of 1:10:000 and are currently the most accurate reference data for the study area. The geometric precision of the layers is less than or equal to 5 m. The minimum mapping unit is 200 m^2 for agricultural and natural areas, while for artificial areas, water bodies, wetlands, rivers and riparian vegetation is 50 m^2. The minimum width for linear elements is 10 m except for the road and rail network which have no limitation. The error rate is between 1.5 and 3%, depending on the class, according to the analysis of 200,000 polygons, which represents 20% of the surface area of Andalucia (87,268 km^2) (Gil *et al*., 2010). The LULC inventories cover years 2005, 2009, 2011 and 2013, being the last available the one used in this research since it is the best reference available at the date of the study.

The LULC classifications of LT8 and S2 imagery were produced by object-based classification using GeoClassifier, a GIS software developed by Jena Optronik in the framework of SWOS and which is used in the project for LULC mapping tasks. GeoClassifier follows a classic workflow of this kind of classification (Blaschke, 2010; Geneletti and Gorte, 2003; Laliberte *et al*., 2004). It begins with a segmentation to define the objects from the satellite image, continues with a training step of the segments where LULC classes are assigned to them, and ends with the calculation of statistics and the final classification of the whole image.

The segmentation is controlled by two parameters: tolerance and minimum mapping unit (MMU). The tolerance defines the range of the pixel value differences in a segment that is used to decide if a pixel belongs to a segment or not, while, on the other hand, the MMU controls the size of the smallest segment that will be generated in the process.

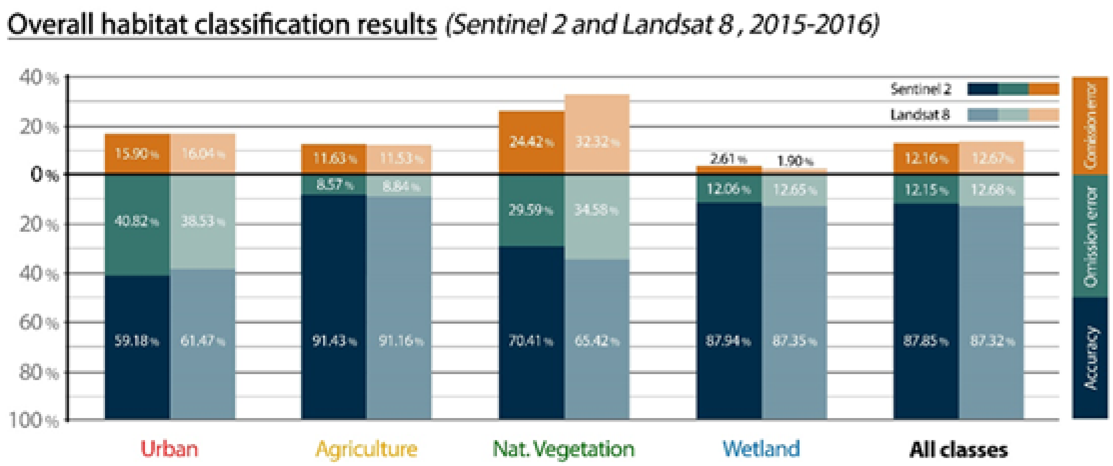
The training process was supported by aerial photographs of recent dates provided by Junta de Andalucia and images available through Google Earth. To this was added a good local knowledge of the area from previous studies which helps to improve the classification process and promotes the quality of the results (Lu *et al*., 2010; Manandhar *et al*., 2009).

Finally, GeoClassifier performs a maximum-likelihood classification, a method in which unknown pixels or objects are assigned to classes using contours of probability around training areas using the maximum-likelihood statistic (Allaby, 2008). All segments are assigned to the class with the highest likelihood. GeoClassifier also includes a tool to make quickly post classification corrections, so that the researcher can change the class assigned to a specific segment manually. However, in order not to introduce variables that may hinder the interpretation of the results, this functionality has only been used in large segments where the true LULC was known accurately.

After the LULC classification, an accuracy assessment was carried out using the error matrices with the 12 LULC classes. All the polygons of the reference map were overlapped with both LT8 and S2 classified maps to produce the error matrix and calculate the accuracy and reliability, also known as producer's and user's accuracy respectively (Congalton and Green, 1999; Congalton, 1991), as well as the KAPPA coefficient (Cohen, 1960). Moreover, the coefficient of determination (R^2) and the root-mean square error (RMSE) were calculated as another agreement measure between the surface detected for each LULC class using the satellite images and those reported in the inventories.

Results

Fig. 3.36 present overall accuracy assessment results for the four habitat groups (urban, agriculture, natural vegetation and wetland) which represent the sum of the surfaces of the 12 LULC classes. The column graph shows accuracy and omission error values as percentage in the lower part of the ordinate axis and commission error in the upper part, so that the magnitude of both errors is represented together. Under the graph, the areas reported by the reference maps (IN - Inventory) and those obtained from the classification of the images are shown. These are presented in four classes: 1) Detected (D), which is the total area of an LULC class on the classified map; 2) Accurately detected (AD), understood as the pixels that have been correctly classified, that is, those that match the inventory. It would be the surface related to the accuracy; 3) Not detected (ND), which corresponds to the omission error area; and 4) Overestimated (O) or the area belonging to the commission error.



	Urban Area (km²)		Agriculture Area		Nat. Vegetation Area		Wetland Area		All classes Area	
IN	10.46		157.17		15.24		13.60		196.47	
	S2	LT8	S2	LT8	S2	LT8	S2	LT8	S2	LT8
D	7.36	7.67	162.62	161.96	14.21	14.73	12.28	12.10	196.47	196.47
AD	6.19	6.44	143.71	143.28	10.73	9.97	11.96	11.88	172.59	171.56
ND	4.27	4.03	13.47	13.89	4.51	5.27	1.64	1.72	23.88	24.90
O	1.17	1.23	18.91	18.68	3.48	4.76	0.32	0.22	23.88	24.91

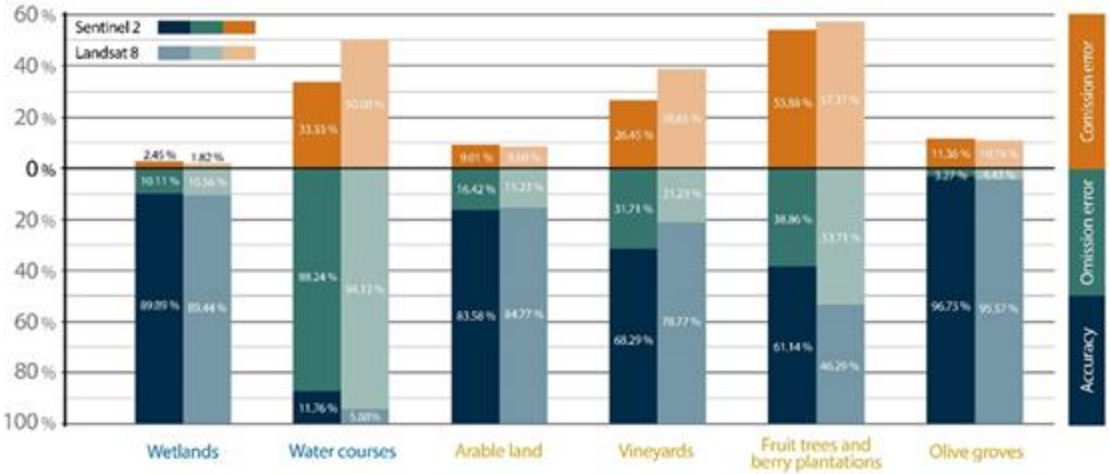
IN - Inventory D - Detected AD - Accurately Detected (Accuracy) ND - Not Detected (Om. Error) O - Overestimated (Com. Error)

Figure 3.36: Summary of Landsat 8 and Sentinel-2 classification results in area (km2) and percentage: overall and LULC groups.

The segment-based classifications provided reliable and accurate results using both satellites. Overall accuracy and reliability is around 88-87% in both cases being slightly better in S2 with an improvement of approximately, 0.5% which is a negligible difference. The comparison of S2 and LT8 error matrices shows that they are not significantly different. The R^2 obtained in the linear regressions between the LULC inventory surfaces and the classified maps are also similar in both satellites, 0.994 for S2 and 0.995 for LT8, and the same happens with the NRMSE, 0.028 and 0.024 respectively. Therefore, using a standard classification method, S2 and LT8 seem to offer very similar results.

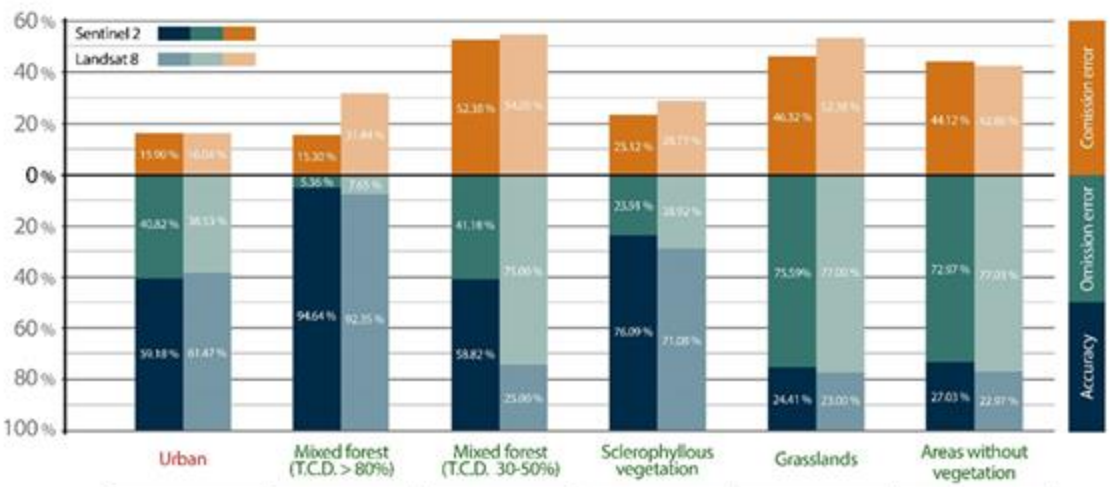
The results of each of the 12 LULC classes, summarized in Fig. 3.37, show more similarities and differences between both classified maps. Wetland class presents very high accuracy values, close to 90%, with a very low commission error. It can be said that both satellites give a good result for this class. However, it must be considered that the wetlands in the area are mostly water bodies, which are not usually difficult to detect with optical imagery. The areas with vegetation of these wetlands are of scarce surface in comparison and nevertheless it is where most of the error of omission originates, especially where the density of vegetation is low because the signal of the bare soil seems to produce confusions with classes of agriculture, urban and other natural vegetation of low density. On the other hand, water courses have not been correctly classified and have an omission error level above 88% in S2 and 94% in LT8, with high commission errors too. This was expected since this class is composed of small and narrow streams with none or very low flow rates for most of the year. Therefore, they are elements that are beyond the resolution of both satellites. S2 detects almost three times more surface enough to make a confident classification of this class.

Habitat classification results *(Sentinel 2 and Landsat 8, 2015-2016)*



	Area (km²)		Area		Area		Area		Area		Area	
IN	13.26		0.34		50.17		3.91		1.75		101.35	
	S2	LT8	S2	LT8	S2	LT8	S2	LT8	S2	LT8	S2	LT8
D	12.22	12.08	0.06	0.02	46.08	46.52	3.63	5.02	2.32	1.90	110.59	108.52
AD	11.92	11.86	0.04	0.01	41.93	42.52	2.67	3.08	1.07	0.81	98.04	96.86
ND	1.34	0.22	0.02	0.32	8.24	7.64	1.24	0.83	0.68	0.94	3.31	4.49
O	0.30	1.40	0.30	0.01	4.15	4.00	0.96	1.94	1.25	1.09	12.55	11.65

IN - Inventory D - Detected AD - Accurately Detected (Accuracy) ND - Not Detected (Om. Error) O - Overestimated (Com. Error)



	Area (km²)		Area		Area		Area		Area		Area	
IN	10.46		3.92		0.68		7.78		2.13		0.74	
	S2	LT8	S2	LT8	S2	LT8	S2	LT8	S2	LT8	S2	LT8
D	7.36	7.67	4.38	5.28	0.84	0.37	7.70	7.75	0.95	1.05	0.34	0.28
AD	6.19	6.44	3.71	3.62	0.40	0.17	5.92	5.52	0.51	0.49	0.19	0.16
ND	4.27	4.03	0.21	0.30	0.28	0.51	1.86	2.35	1.61	1.64	0.54	0.57
O	1.17	1.23	0.67	1.66	0.44	0.20	1.78	2.23	0.44	0.56	0.15	0.12

IN - Inventory D - Detected AD - Accurately Detected (Accuracy) ND - Not Detected (Om. Error) O - Overestimated (Com. Error)

Fig 3.37: Landsat 8 and Sentinel-2 classification results by LULC class.

The visual comparison of the results, presented in Fig. 3.38, show that the maps generated from the S2 and LT8 imagery are very similar to each other, and they have to a large extent the same information as the LULC map used as a reference.

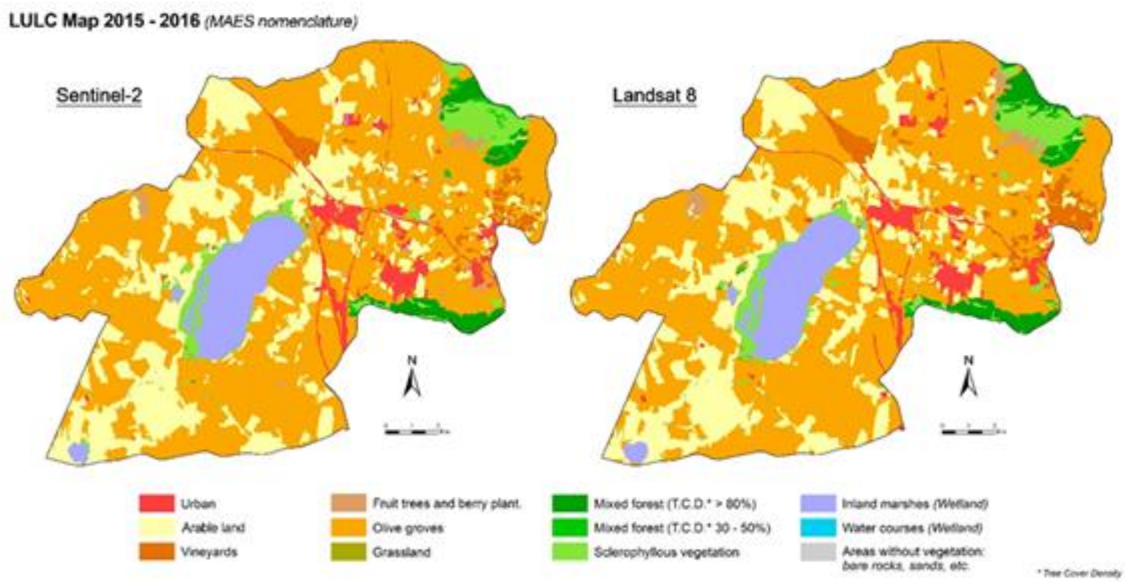


Fig 3.38: LULC maps produced from the Landsat 9 and Sentinel-2 classification.

The results obtained for the classification of S2 and LT8 imagery were very similar. Both satellites offer good accuracy and reliability for general use in the LULC monitoring of Mediterranean wetland areas. S2 showed some improvements thanks to its better spatial and spectral resolution but it is not a remarkable difference as to ignore or to replace the use of LT8. Therefore, both satellites can and should coexist and be used together to increase data availability providing one the lack of images of the other in order to have the best possible results in remote sensing researches.

Monitoring surface water dynamics

Material and methods

The Surface Water Dynamic (SWD) indicator derived from Landsat imagery for January 2007 to September 2015 was calculated by using the Modified Normalized Difference Water Index (MNDWI; Hanqiu Xu, 2006). MNDWI was calculated by the spectral responds of Green and Short Wave Infra-Red (SWIR) bands of Landsat. The MNDWI identify/discriminate by raster cells the presence or absence of water. The resulting raster values have a range from -1 to +1, where negative values indicate no water content raster cells and positive values indicate water presence. 0 to +1 values were extracted from the resulting layers to produce a binary water mask where value 1 represents water and value 0 no water. To produce the SWD TF product, all these masks are summed and the resulting layer is then normalized from 0 to 1 or 0 to 100 to express the flood periods in percentage (Fig. 3.39). For the assessment approach, the masks were kept separately and the surface water extent of each one was calculated in hectares (the total area that the positive MNDWI cells cover). This surface is understood as open surface water extent area (SWE).

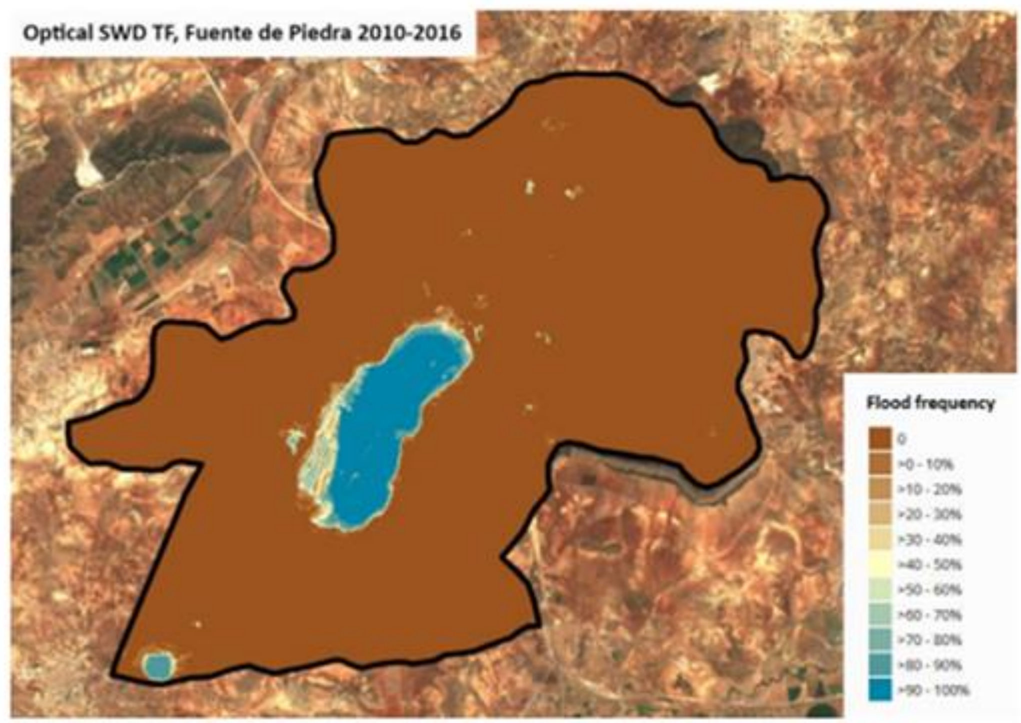


Fig 3.39: Example of optical SWD TF product of Fuente de Piedra. Years 2010 to 2016.

The assessment of the studied was done using as in-situ data the environmental variables temperature, precipitation and water level in the wetland. These data were obtained from the *Instituto Geologico y Minero de España* (IGME, 2000) and the Nature Reserve authorities of Fuente de Piedra wetland for the hydrological period January 2007-December 2017. Monthly average temperature and precipitation were calculated from two different meteorological stations located within the wetland basin (Cero del Palo, Herriza-Fuente de Piedra). The piezometry data are monthly and daily level measurements obtained from a limnigraph located in a well in almost the centric part of the lagoon. Assessment for the optical SWD product was made by observing the statistical correlation between the monthly SWE and the monthly measurements of the selected environmental variables (precipitation, temperature, water level). The risen hypotheses were: 1) Is there statistical correlation between temporal changes in SWE and temporal changes in piezometry level, precipitation and temperature? 2) Which is the most correlated variable with the SWE? 3) What is the confidence level of this assessment?

Statistical analyses of correlation and linear multi regression were conducted. We applied the Spearman correlation analyses (variables were not normally distributed); respectively how the SWE varies with the water level, precipitation and temperature variability. Linear multi regression analysis is a cause-effect relationship and it is used to fit a relationship between two variables such that one can be predicted from the other, respectively cause-effect relationship of SWE variable and environmental condition variables is tested.

The assessment process for SAR SWD data from Sentinel-1 (that acquires imagery in C-Band wavelength and mostly in VV polarization and with a temporal resolution of 5 days) was done in comparison with the main environmental variable identified, being the water level. In this case, monthly and daily measurements where used. The purpose was to compare both products and to give answer to the hypothesis: Is the SWD indicator derived by SAR increasing accuracy once compared to Landsat SWD indicator?

Spearman correlation analyses (variables were not normally distribution) and linear multi regression analysis were applied for two temporal distribution variables: a) Mean SWE SAR (in ha/month) and mean water level (in cm/month) and b) Daily SWE SAR (in ha/day) and daily water level (in cm/day).

<u>Results</u>

In the case of Optical SWD assessment, statistical correlations results show a significant positive correlation (0.932) between SWE and water level. Temporal trends in SWE and water level move in the same direction, having a slight rise from 2007 to 2015, once precipitation trendlines seems to slightly decrease from January 2007 to September 2015 (Figure 3.40). However, hydro-graphical results show clear correlation between the water level, SWE and the precipitation during extremely wet years where the high precipitations can be remarkable in water level and SWE but slightly displaced in time (usually in one-month displacement). For example, high precipitation during December 2009 (230mm) and 2010 (140mm) maintain the water level in maximum fee/quotas (150-170cm) and the SWE in maximum extent (1200-1400ha) (blue circles, Figure 3.40).

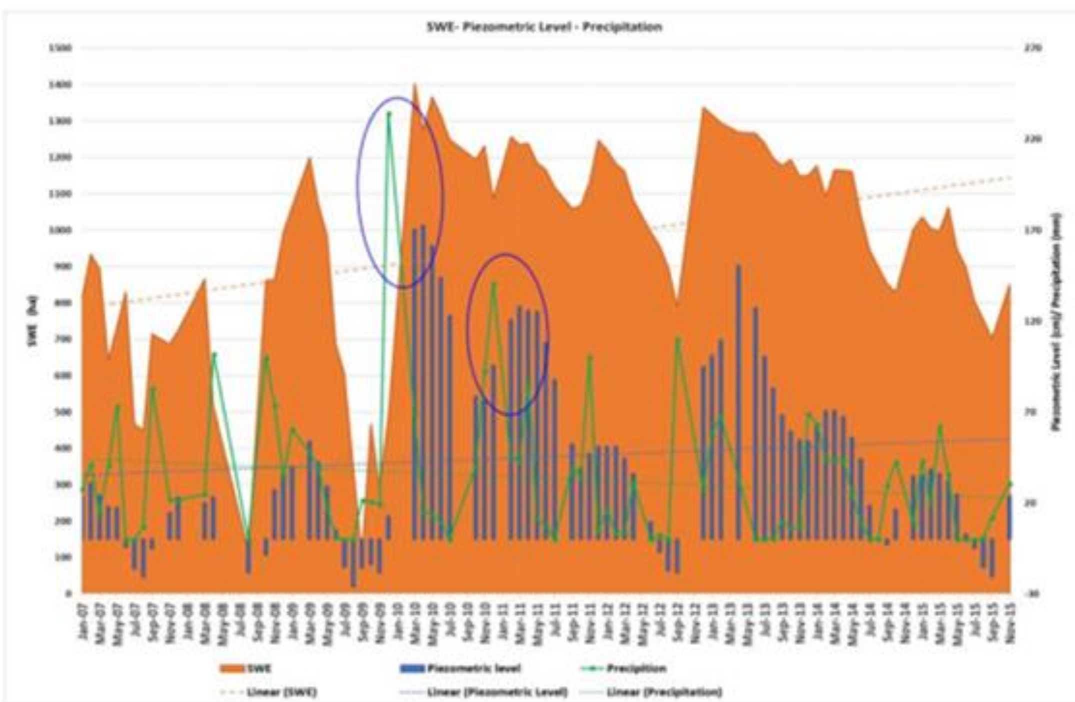


Fig. 3.40: Trend lines between SWE, water level and precipitation variability for the period January 2007-September 2015. Blue circles indicate the highly precipitation months and the notable displaced in the next months.

Further, SWE variable show no significant correlation with precipitation variable and significant negative correlation with temperature variable (Table 3.8). No significant correlation with precipitation variable might be explained by other hydro-geological effects in the endorheic basin that were not considered in this assessment analysis. Thus, we can expect that not all occurred precipitations are reflected in the wetland SWE. Besides, in the last years it has been observed an important increase in the groundwater pumping mainly by agriculture exploitation, which have decreased the water level in the basin and reflected in the SWE level. As about the negative correlation with temperature variable, it is of clear understanding that once temperature rise the SWE decries and vice versa. However, to have a better understanding in SWE and temperature correlations it would be necessary to take in account other variables as evaporation, evapotranspiration etc. which are beyond the objectives of the assessment methodology described in this report.

Table 3.8: Correlation results

		Mean Temperature (°C)	Mean Precipitation (mm)	Water level (m)
SWE (ha)	Correlation Coefficient	-,296**	,118	**,932****
	Sig. (2-tailed)	,006	,279	,000
	N	86	86	86
Mean Precipition (mm)	Correlation Coefficient	-,597**	1,000	,215*
	Sig. (2-tailed)	,000	-	,047
	N	86	86	86
Piezometriv Level (m)	Correlation Coefficient	-,351**	,215*	1
	Sig. (2-tailed)	,001	,047	-
	N	86	86	86
Mean Temperature (°C)	Correlation Coefficient	1,000	-,597**	-,351**
	Sig. (2-tailed)	-	,000	,001
	N	86	86	86

Results from the multi regression model shows a $R^2 = 0.644$ and water level as the first important predictor for the SWE, while both temperature and precipitation predictor variables were excluded (Table 3.9). This means that SWE is the best variable in predicting the water extent.

Table 3.9: Multi regression analysis results.

Model	R	R Square	R Adjusted Square	Std. Error of Estimate	Durbin-Watson
1.00	,803a	,644	,640	170,890.00	,683

a. Predictors: (Constant), Water level (cm)

High predictive capacity by water level ground data allows to generate precise simulation models that can help in monitoring water level fluctuations, flood regime and water availability. However, the total predictive level of regression model raises uncertainties. This might come regarded to SWE indicator limitations: 1) the SWE has its limitation in detecting water content below vegetation cover, or soils highly saturated in water content; 2) the SWE extraction date does not overlap with the in-situ water level measured date; 3) comparted monthly mean water level, precipitation and temperature with SWE of precise day of same month might create disparity.

As for the SAR SWD assessment, Spearman correlation analysis shows high correlation factor of 0.936 between mean SWE SAR and mean piezometry level (a). However, multi regression analysis shows that the strength of the predictive model is quite higher (R^2= 0,849; Table 3.9) while compared to Landsat (R^2= 0.644; Table 3.9), which means that accuracy of predicting SWE by SAR is much higher than predicting from Landsat.

Table 3.10: Multi regression analysis results.

Model	R	R Square	R Adjusted Square	Std. Error of Estimate
1.00	,920[a]	,847	,841	14,037,534.00

a. Predictors: (Constant), Mean Water level (cm)

The second assessment process using the daily variables (b) shows not significant differences in correlation factor, unlike the correlation between daily SWE_SAR and waterlevel seems to be slightly lower (0.858) compared to Landsat SWE (0. 932), explained by the fact that additional hydrological behaviours as the surface evaporation and evapotranspiration processes, especially during the summer, are translated into underground flows and thus in the water level. However, Fig. 3.41 shows clearly how the daily trends in water level are reflected in the SWE. Approaching in the dry summer period, it is observed a short transit period where the water level remains in positive values while the SWE is almost very low or 0 (red circle Fig. 3.41).

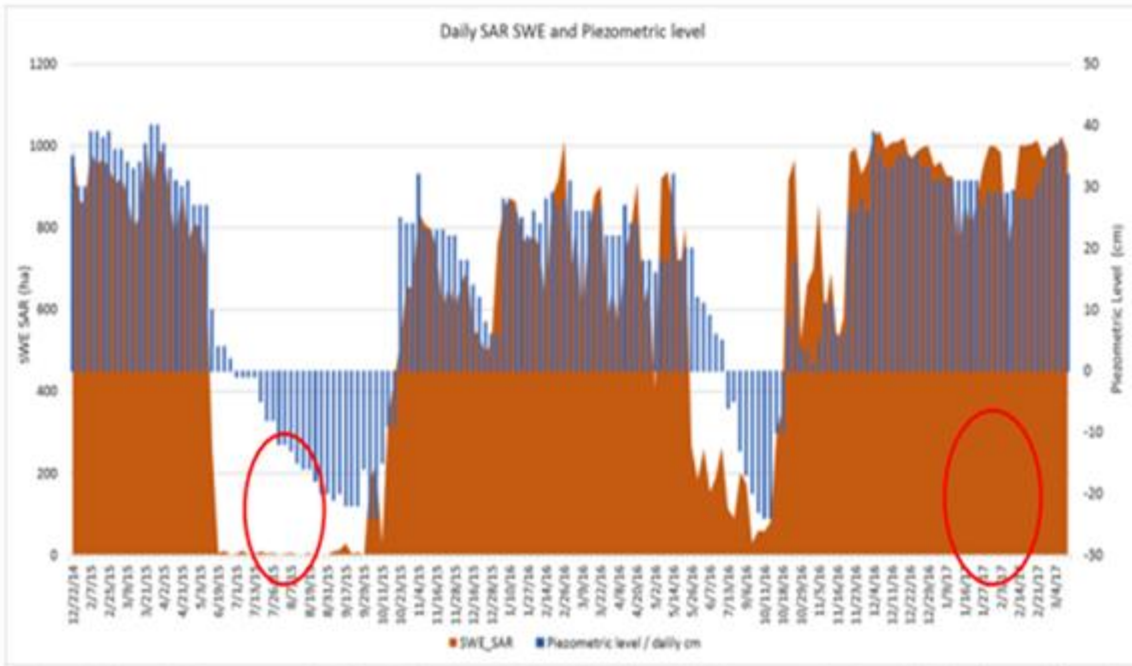


Fig 3.41: Graphic between daily SWE-SAR variability (in ha) and daily water level variability (in cm) for the period December 2014- March 2017.

The overall result indicates the high capacity of both SWD products to identify very low water presence in the land surface and give accurate estimations of the water quantity we have available for corresponding water level and its variation.

SAR SWD indicator seams to give high predictive capacity comparing the use of Landsat SWD indicator. However, it would be proper to have a longer assessment period to minimize the level of uncertainty, since Sentinel 1 data are only available from 2014.

Once compare to SAR SWD indicators, the optical SWD indicator seems to overestimate the SWE detected. This probably comes from the low temporal resolution of Landsat imagery (14 days), where SWE were calculated as a mean of maximum three images per month. In case of SAR, the SWE were calculated as a mean of maximum eight images per month, which increase the detection on monthly SWE variation (temporal resolution of Sentinel 1 is 5 days, and without cloud limitations).

As conclusion, we could affirm that SWD indicator (by both sensor Landsat, SAR) is a proper indicator that can help in monitoring water dynamics, identifying areas with high water accumulation in high runoff periods or areas that remain water during extreme drought periods. Those, information supports stakeholders in planning purposes in terms of mitigation, areas of immediate action, areas of strict conservations, etc.

Bibliography

Abdul Malak, D., Sánchez-Espinosa, A., Schröder, C. (2016). Guidelines for the delimitation of wetland ecosystems (v.1.1). SWOS Technical publication.

Allaby, M., 2008. A Dictionary of Earth Sciences, 3 ed. Oxford University Press, Oxford. https://doi.org/10.1093/acref/9780199211944.001.0001. 13: 9780199211944.

Blaschke, T. (2010). Object based image analysis for remote sensing. ISPRS journal of photogrammetry and remote sensing, 65(1), 2-16.

Carnell, P. E., Windecker, S. M., Brenker, M., Baldock, J., Masque, P., Brunt, K., & Macreadie, P. I. (2018). Carbon stocks, sequestration, and emissions of wetlands in south eastern Australia. Global change biology, 24(9), 4173-4184.

Catalán, F., Linares, L., López-Geta, J.A., Rubio, J.C., Rodriguez Padilla, J.L., Fernández del Río, G. et al. (2002). Homenaje a Manuel del Valle Cardenete. El sistema hidrico de la Laguna de Fuente de Piedra, situación actual y perspectivas futuras. Instituto Geológico y Minero de España (IGME), Madrid, Spain.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), 37-46.

Congalton, R.G., 1991. A review of assessing the accuracy of classification of remotely sensed data. Remote Sens. Environ. 37, 35–46.

Congalton, R.G., Green, K., 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers, Boca Raton, USA.

Geneletti, D., & Gorte, B. G. H. (2003). A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. International Journal of Remote Sensing, 24(6), 1273-1286.

Gil, Y., Ortega, E., Dominguez, M. C., Navas, P., Patiño, M., Vicent, C., ... & Moreira, J. M. (2010). SIOSE Andalucía, experiencia de integración y actualización de bases cartográficas multiescala. In Congreso Nacional de Tecnologías de la Información Geográfica (14. 2010. Sevilla)(2010), p 116-134. Universidad de Sevilla.

González Báez, R., Gavilán, J., and Vadillo Pérez, I. (2017). Análisis de tendencia de contaminantes en la masa de agua subterránea Fuente de Piedra (provincia de Málaga). 63 Sesión Científica Sociedad Geológica de España. Elche, November 2017; uri: hdl.handle.net/10630/14788.

Grasset, C., Rodriguez, C., Delolme, C., Marmonier, P., & Bornette, G. (2017). Can soil organic carbon fractions be used as functional indicators of wetlands?. Wetlands, 37(6), 1195-1205.

Hanqiu, Xu. (2006). Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. International Journal of Remote Sensing, 27(14): 3025–3033.

IGME (2002): Instituto Geologico y Minero de España, "3 : Climatologia". Available at: http://www.igme.es/zonas_humedas/laguna/medio_fisico/ev_piezo_lib99/pdf/lib99/in_03.pdf

Laliberte, A. S., Rango, A., Havstad, K. M., Paris, J. F., Beck, R. F., McNeely, R., & Gonzalez, A. L. (2004). Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. Remote sensing of Environment, 93(1-2), 198-210.

Lu, Q., Chen, Q., Hou, X., Bi, X., & Zhao, Z. (2010, October). Using local knowledge to map crop types and other land use/cover types with a limited number of TM images in Shandong Peninsula, China. In 2010 3rd International Congress on Image and Signal Processing (Vol. 5, pp. 2287-2293). IEEE.

Manandhar, R., Odeh, I. O., & Ancev, T. (2009). Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. Remote Sensing, 1(3), 330-344.

Martos-Rosillo, S., González Ramón, A., Heredia Díaz, J.G., Pedrera Parias, A., and Rodríguez Rodríguez, M. (2017). Changes in groundwater quality due to intensive exploitation of the carbonate aquifer of Sierra de Humilladero (Malaga, Spain). Geogaceta, 62:91-94. ISSN: 0213-683X.

Mitsch, W. J., Bernal, B., Nahlik, A. M., Mander, Ü., Zhang, L., Anderson, C. J., ... & Brix, H. (2013). Wetlands, carbon, and climate change. Landscape Ecology, 28(4), 583-597.

Rodriguez-Rodriguez, M., Martos-Rosillo, S., and Pedrera, A. (2016). Hydrogeological behaviour of the Fuente-de-Piedra playa lake and tectonic origin of its basin (Malaga, southern Spain). Journal of Hydrology, 543B:462-276; doi: 10.1016/j.jhydrol.2016.10.021.

Strachan, I. B., Nugent, K. A., Crombie, S., & Bonneville, M. C. (2015). Carbon dioxide and methane exchange at a cool-temperate freshwater marsh. Environmental Research Letters, 10(6), 065006.

Warwick, N. W., & Brock, M. A. (2003). Plant reproduction in temporary wetlands: the effects of seasonal timing, depth, and duration of flooding. Aquatic Botany, 77(2), 153-167.

Success story #2 for LULUCF report in Spain. Förecast: estimation of extracted biomass in forestland.

Keywords: biomass, changes, Sentinel, LiDAR, algorithms, LULUCF

Application field: estimation of biomass extracted in clearcuts.

Abstract

Forests play a dual role in their relationship with Climate Change: on the on hand, forests suffer the negative effects of climate change (forest fires, pests, and other big disturbances) but, on the other hand, are part of the solution. Forest are huge carbon sinks, and their sustainable management is key to maximize this carbon uptake from forests. In a first moment, as trees absorb carbon through photosynthesis but, second, and not less important, through the sequestration in forest products. Some of these products act as long-term sequestrators of the carbon absorbed by trees, i.e., wooden structures. Others, although having a very short life cycle, are very important as they replace other materials. Among these ones, biomass outstands as the main substitute for fossil fuels. Thus, estimation of biomass extraction appears as a key element in a LULUCF framework.

Biomass represents an increasingly important player in the forest value chain, not only from a LULUCF and Climate Change perspective (as said, as replacement for fossil fuels), but also represents a big part of the wood industry, with the installation of hundreds of biomass processing industries in Europe in the last decade. The use of biomass for heating is in the agenda of most of the public bodies in Europe and the information on available supply for the industry is necessary. In the near future, this need will increase almost exponentially, so offering decision support systems on the availability (location, quantity, etc.) of biomass for the industry and, subsequently, the final users will be key.

In this example, we use a two-step methodology: the estimation of biomass uses the combination of satellite and LiDAR information in order to detect clearcuts and, subsequently, quantify the biomass extracted. This shows an example of how the estimation of extracted biomass can be easily assessed for future needs.

Methodology used / Workflow

The methodology used is comprised in 'förecast', a forest intelligence platform. Specifically, in order to estimate biomass extraction, two different methodologies are used. As a first step, a change detection methodology is used in order to detect where and when a particular area has suffered changes (usually, clear cuts). This first methodology is based on vegetation index calculation from satellite information (Sentinel-2) and, through Artificial Intelligence methods, the system is fed with recurrent information in order to adjust the level of change or the type of change detected. The main vegetation indexes that are automatically estimated in forecast are NDVI (Normalized Difference Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), SAVI (Soil Adjusted Vegetation Index) and CI GREEN (green chlorophyll index), but any other interesting Vegetation Index can be calculated.

Using NDVI as a basic index for changes, an iterative work of change rate in NDVI has been analyzed in order to reach the optimal levels of change in every ecosystem that maximizes the change detection while minimizing false positives. Machine learning and other robust statistical analysis are developed in order to obtain this optimal automatic result.

The second step involves LiDAR information and the development of biomass algorithms for the present species in the areas detected as clear cuts. First, the LiDAR point cloud is processed to extract the metrics in the study area. Then, using already developed models or field information, a biomass algorithm is created based on LiDAR metrics. This algorithm is then applied to the estimated biomass extracted area.

For the example of the solution, a small Area of Interest (AOI) of around 300ha was used. The AOI is located in Espirra, close to Lisbon, in Portugal. For that reason the SRC of the cartography is ETRS89 / UTM zone 29N (EPSG 25829). This example was chosen due to its active management, so the detection of changes is clear and the change rate is very high compared to the mean forest in Europe. The AOI is a *Eucalyptus globulus* plantation and this is the reason of the change rate. Also, in areas like this one, with very productive forests, is where the change detection and the calculation of biomass is really important.

Fig. 3.42: Location of the AOI in Espirra, Portugal

Figures 3.43 and 3.44 show how the changes in the AOI are detected through förecast, a platform allowing the automatic calculation of different Vegetation Indexes. Among them NDVI is one of the main Vegetation Indexes used, so the example shown here is based on this Vegetation Index.

Figure 3.43 shows the calculation of NDVI in a particular date (in this case, 16/11/2021). This figure shows how NDVI reacts to the clearcuts in the northeastern part of the AOI, where greener colors represent high NDVI values (high vegetation activity) while yellow-brownish colors represent low NDVIs. It also shows the date selection tool, where all the available Sentinel-2 images are automatically processed and Vegetation Indexes calculated.

Fig 3.43: NDVI estimation in förecast for the AOI (16/11/2019)

Finally, Fig. 3.44 shows how förecast's change detection system works. Based on NDVI, the platform shows the evolution of the vegetation since mid-2015 for each of the management units of the AOI. In this evolution, the green lines show growth or equal levels of NDVI, while orange lines show small changes, and the red lines show big changes in NDVI. Usually, the red lines correspond with important changes in vegetation (clearcuts, fires, etc.) and this change detected in förecast comes with a service alert via email or, if need, through an API to connect to the client's systems.

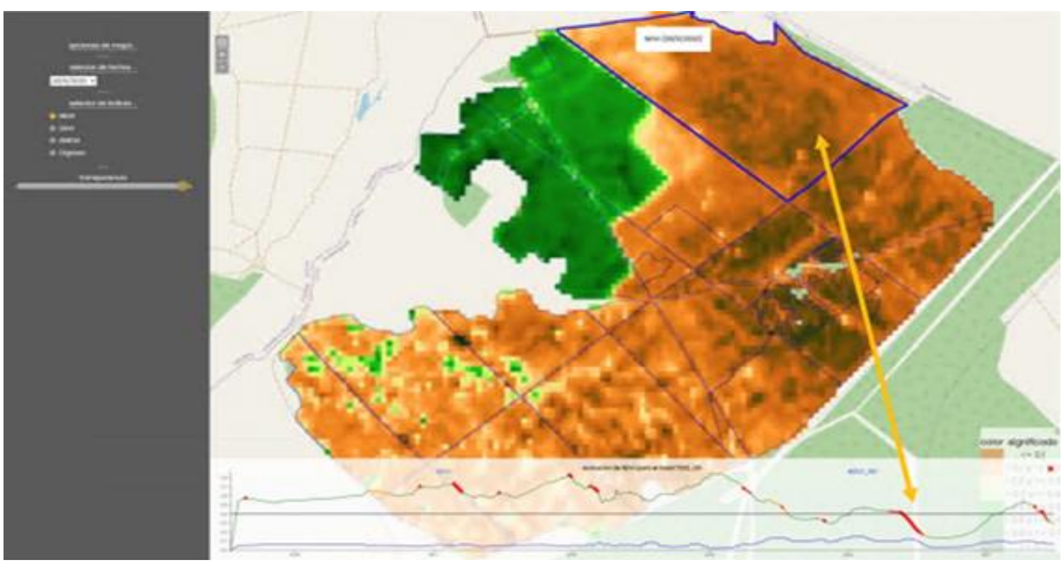


Fig. 3.44: Change detection system in förecast shown the evolution of NDVI and how a clearcut (NE) is detected in the NDVI line.

Once the change (clearcut) is detected, the algorithms created for the species and the area allow us to know the quantity of biomass extracted. Through the same platform, fӧrecast, a calculation of biomass per management unit is first visualized (Figure 3.45) and then, it can be downloaded (Figure 3.46) so the results by pixel (usually 20x20m pixels) can be shown in any GIS (Figure 3.47). This information is key because through simple geostatistical procedures, the total biomass extracted can be calculated, or the biomass extracted in a particular area of the AOI, etc.

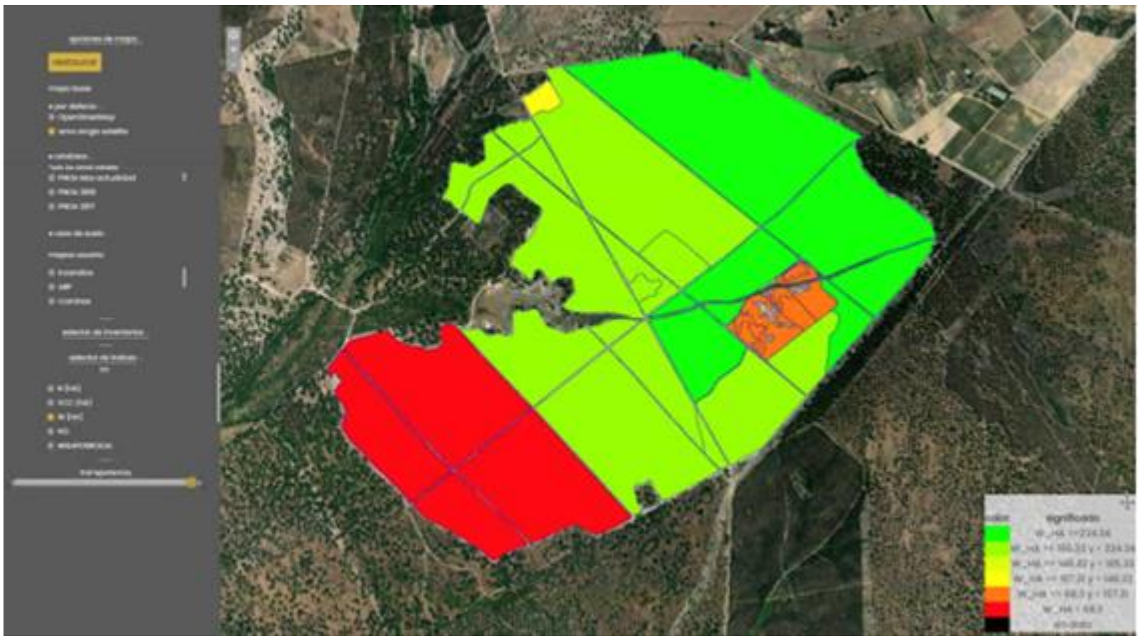


Fig 3.45: Visualization of the biomass extracted by management unit in forecast.

Fig 3.46: Download interface in forecast.

The most important limitation of this system is the need of a LiDAR flight in order to be able to develop the algorithms that will be applied and will give the results on biomass extracted (or any other forest inventory variable). In this case, an ad hoc LiDAR flight was done in june 2019. When an ad hoc flight is not possible, other LiDAR sources may be used (i.e. the PNOA flights in Spain).

Fig. 3.47: Representation of values of biomass extracted in the AOI in QGis from downloaded information of förecast.

<u>Key results</u>

Change detection as a key factor for LULUCF and estimation of biomass extracted in the areas detected as clear cuts. With all that, a clear estimation of biomass extracted and its related information (carbon absorbed by the forest, fossil fuels replaced, etc.). The results are spatially continuous, allowing to have information for any area needed. The results can be obtained by management unit (as shown in Figure 3.45) and also by 20x20m pixels, where every cell has information of biomass extracted (by hectare). The aggregation will give the total biomass extracted for any area (or the total) of the Area of Interest.

Innovative impact

The solution presented offers a clear innovation through the combined use of Copernicus satellite information (Sentinel 2) and LiDAR and the development of algorithms. This combination allows to have almost real-time information on the biomass extracted, thanks to the 5– 7-day refresh from Sentinel and the change detection algorithms created and also due to the application of LiDAR based algorithms in the detected areas. So, this solution shows the potential of Copernicus and LiDAR for real-time and precise geoinformation as a basis for a decision support system for LULUCF standards.

Related Copernicus domains: Forestry, Climate Change and Environment, Energy and Natural Resources.

Area of interest: The Area of interest for this example is a 300ha forest in Espirra, Portugal. It is a Eucalyptus plantation with a high interest in productivity.

Algorithm and data used: Copernicus Sentinel-2; LiDAR information (drone-based ad hoc flight); field information. Algorithms: Change detection and biomass estimation algorithms (self-developed). Change detection algorithms are usually based on NDVI as vegetation index.

Success story #3 for LULUCF report in Spain. Mapping woodland and forests through remote sensing and AI: towards forest species identification in afforestation/reforestation areas

Keywords: Forest classes/ Reforestation/Afforestation/LCLU/ LULUCF

Application field: Mapping LULUCF classes

Abstract

Forests management practices could be crucial to get an accurate value of carbon sinks (Erb *et al*., 2013). Additionally, several studies suggest that the amount of soil organic carbon depends on tree species composition (García-Oliva and Masera, 2004). Thus, mapping forest classes in afforestations and reforestations is a priority in order to effectively evaluate and calculate carbon stocks dynamics in the LULUCF sector, nowadays identified as a Gap in knowledge for LULUCF Regulation.

COTESA presents as a possible solution to this Gap the successful case of the algorithm LMAPP, developed within the R&D project “TIERRA 3” for monitoring land dynamics, for the Xunta de Galicia. TIERRA 3 is a BigData and CloudComputing processing platform for the automatic mapping of Land Cover. The algorithm LMAPP, core of the automated mapping of Land Cover (LC) classes of TIERRA 3; is based on segmentation and classification strategies of advanced remote sensing techniques and supported by artificial intelligence. As input data, time-series of HR multispectral images (Sentinel 2) or VHR multispectral images (satellite or from UAV) could be used. As a result, a LC map is obtained based on the SIOSE (Spanish Land Cover / Land Use Information System) classification, but with unique classes and a minimum mapping unit of 300m2. The platform includes a set of customized edition tools that allows users to review the automatic results in order to get the VHR LC map product SIOSE 5K.

Our research concludes that the best method was the use of two VHR satellite images (one of spring and other of winter within the same natural year) and specific samples extracted for each zone and set of imagery, obtaining average thematic and geometric accuracy of 75%. The accuracy of the forest classes (coniferous forest, deciduous broadleaved forests and evergreen broadleaved forests) was around 85%.

Currently, the algorithm LMAPP is applied in combination with change detection methods to detect forest and tree changes and, as a global LC mapping method; is being tested in other regions of Spain. The results of these new applications suggest that the algorithm LMAPP is a good candidate to improve the mapping of afforestation and reforestation areas by identifying, at least, the forest type. Moreover, the next research line will be focus on the improvement of the algorithm in case of other sources (specifically, with Sentinel -2) and test its functionality at tree species level. Related to that, COTESA intends in the long term to investigate the potential of hyperspectral data and its applicability through spectral resampling techniques to satellite imagery.

Our algorithm constitutes a promising starting point to develop a solution for this Gap, currently providing automated forest types mapping. It is important to highlight that this classification includes a new level in the case of broadleaved forests, in comparison with current Copernicus LCLU map products (CLC, Forest HRL or Local Land Monitoring products): our algorithm detects both deciduous and evergreen of broadleaved forests, key in Mediterranean areas, improving the thematic accuracy of Forest Layer in LULUCF classes.

The strategy proposed by COTESA in order to apply the solution to a national or European scale is based on the use of Sentinel-2 due to its extensive coverage and high Near-Real Time latency. Thus, afforestation and reforestation areas would be detected in order to determine the tree forest composition. Regarding this purpose, it is important to note that accuracy (both thematic and geometric) will be limited to the resolution of the input data (10 m). The success of automate forest species at these scales using Sentinel-2 should study in the future other methodological approaches. In that regard, COTESA presents our recent research in super-resolution imaging: the algorithm IRIX4, which obtains 2.5m resolution images from Sentinel maintaining the spectral data.

In summary, our algorithm LMAPP is presented as a solution to the gap of identification of forest types in afforestation and reforestation areas by automated mapping of coniferous, deciduous broadleaved and evergreen broadleaved tree covers from VHR imagery. Moreover, currently LMAPP is being improved, specialized and tested in different regions of Spain and by means of Sentinel-2 as input data, with promising results. Future researches will evaluate the support and applicability of hyperspectral data and super-resolution imaging techniques to deal with the challenge of getting the tree species level and national and European scales.

Methodology used/ Workflow

LMAPP algorithm: development and methodological approach

The algorithm LMAPP was developed as the core the R&D project "TIERRA 3" for monitoring land dynamics, for the Xunta de Galicia. The main goal was to automate Land Cover mapping according to SIOSE data model in

Galicia (NW of Spain), by remote sensing techniques supported by Machine Learning. The objective was to obtain global LC map products, including urban features, croplands, grassland, shrubs, woodland and forests, sparsely vegetated land, wetland and water bodies' classes.

LMAPP is a pixel-based Random Forest algorithm designed for automating the mapping of SIOSE classes using multispectral data from satellite or UAV imagery. Although LMAPP accepts different input data (mainly Sentinel, WorldView and from UAVs), during the development of the project the solution ended up focusing on 8-bands VHR images due to the high resolution requirements of the final product: minimum mapping unit (MMU) of 300m^2 and 1:5.000 output scale.

The algorithm LMAPP takes into account the spectral data from HR or VHR imagery, and two indices: the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI).

These indices are introduced as bands of the image itself maximizing the spectral response and improving substantially the results, especially of vegetation and water classes:

- The NDVI index makes it possible to identify the presence of green vegetation on the surface and to characterize its spatial distribution, as well as its evolution over time. NDVI values close to 0.1 indicate sparsely vegetated desert areas and values close to 0.9 characterize areas with high vegetation density. It is an index used to estimate the quantity, quality and development of vegetation based on measurement of the radiation intensity of certain bands of the electromagnetic spectrum emitted or reflected by vegetation (Ariza *et al.* 2018).
- NDWI is the most appropriate for mapping bodies of water. The body of water has a high absorption capacity and low radiation in the range of visible wavelengths to infrared. The index uses the green and near-infrared bands of remote sensor images based on this phenomenon. The NDWI can improve water information effectively in most cases. It is sensitive to built-up land and often results in overvalued bodies of water (McFeeters, 1996).

Besides that, our research agree literature (see a review in Xie, Sha and Yu, 2008) that time-series analysis are key to automate the mapping of vegetation classes, so the platform was developed according to the use of two images: ideally, one of spring and one of winter within the same natural year.

A database of the whole territory of Galicia was created to operate as training samples to ingest the algorithm. This database was processed by harmonized integration of several reference databases: SIOSE, Agricultural Plot

Geographic Information System (SIGPAC) and Forest Map of Spain (MFE), among others. However, the different reference years of these databases and the imagery selected for every zone means a factor of error in the results. Thus, specific training samples for every area and set of images resulted as the best option, so the platform was designed in order to operate with new training samples delineated by the users. These specific training samples meet the following requirements:

- Balanced distribution on the whole zone and among samples of the same land cover class.
- Same area proportionality of each class as the reality of the zone.
- Maintain a safety margin at the boundaries and edges of the land cover class in order to avoid errors or noises.

The workflow of the algorithm LMAPP is summarized as follows (Figure 3.48):

1. Sample creation and extraction: random points are generated and selected over the training samples areas, following a density strategy optimized for every class.
2. A multi-temporal statistical analysis of the whole set of multispectral data, NDVI and NDWI (calculated internally in the LMAPP algorithm) is carried out.
3. The algorithm is training based on the sample random points and a customized model.
4. Then pixel-based classifications of the imagery are performed.
5. Finally, finalization processes are implemented in order to get the LC map final product. This processing includes generalization of the result in order to avoid noises and meet the requirements of MMU the final product, also a polygonization and integration of cadastral information.

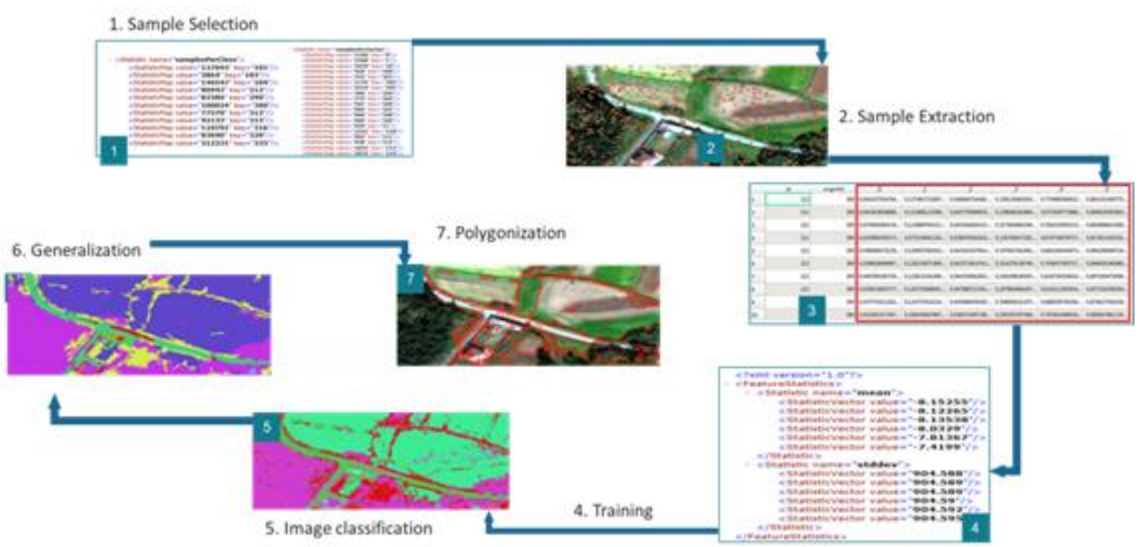


Fig. 3.48: Overview of the workflow of the algorithm LMAPP.

<u>LMAPP algorithm: results</u>

LMAPP was designed in order to automate global Land Cover mapping, i. e. getting a whole set of LC classes following the SIOSE data model but with unique classes: every polygon should represent an only land cover.

During the project, different methodological strategies were performed according to different factors:

- Use of one or two images of the same area.
- Use of different input data in terms of spectral response (number of bands) and resolution.
- Use of the general training sample database or customized samples for every area.

The tests carried out during the development phase of the project suggested that the best scenario was the use of two VHR satellite layers (one of spring and one of winter within the same natural year) and specific training samples.

To check it, a validation phase was set in order to evaluate statistically the results and the platform by tests in different areas than the ones used during the development phase. The validation was performed taking into account the possible combinations of scenarios but with the efforts focused in the best scenario.

The results were validated by photointerpretation, both thematically and geometrically. For the thematic validation, a Stratified Random Sampling was designed for every area tested so that at least 50 points of every class and every area were validated, but adjusting the number of samples taking into account minority, majority and important classes. For the geometric validation, the results were tiled and selected with the focus on land cover transitions inside the territory, covering more than 25% of every area to be reviewed.

The results of the validation of the best scenario are summarized in Table 3.11.

Table 3.11: Validation results of automated classifications in TIERRA 3 with COTESA's algorithm LMAPP, using 2 WorldView images (spring and winter) and customized training samples. 20 different zones in Galicia were automated covering 232 km^2 and validated by photointerpretation in terms of thematic accuracy, meanwhile 17 different zones were validated geometrically, covering 139 km^2.

LMAPP – 2 VHR images and new specific training samples							
SIOSE LAND COVER CLASSES	*THEMATIC ACCURACY*				*GEOMETRIC ACCURACY*		
	Overall Accuracy: 75%		***Kappa: 73%***		***Coincidence***	*Comission*	*Omission*
	User's Accuracy	***Producer's Accuracy***	*Commission Error*	*Omission Error*			
101 - Buildings	**58%**	**72%**	42%	20%	**58%**	22%	20%
102 – Green urban areas	**82%**	**80%**	18%	0%	**0%**	100%	0%
107 – Road network	**75%**	**77%**	25%	17%	**69%**	15%	17%
131 – Extraction and dumping sites	**74%**	**67%**	26%	3%	**66%**	32%	3%
219 – Herbaceous crops and pastures	**87%**	**67%**	13%	8%	**82%**	11%	8%

223 – Non-citrus fruit trees	**41%**	**86%**	59%	0%	**0%**	100%	0%
231 - Vineyards	**56%**	**82%**	44%	27%	**51%**	21%	27%
300 – Natural grasslands	**66%**	**68%**	34%	17%	**63%**	20%	17%
312 – Deciduous broad-leaved forest trees	**85%**	**88%**	15%	6%	**90%**	4%	6%
313 – Evergreen broad-leaved forest trees	**89%**	**75%**	11%	7%	**85%**	8%	7%
316 - Coniferous forest trees	**79%**	**88%**	21%	10%	**85%**	5%	10%
320 - Shrubs	**77%**	**64%**	23%	12%	**73%**	14%	12%
331 – Beaches, dunes and sand plains	**83%**	**83%**	17%	7%	**89%**	4%	7%
333 – Bare soil	**52%**	**81%**	48%	35%	**50%**	15%	35%
352 – Outcrops	**64%**	**91%**	36%	13%	**30%**	57%	13%

421 – Coastal marshes	**76%**	**64%**	24%	5%	**75%**	21%	5%
511 – Water courses	**72%**	**97%**	28%	17%	**81%**	2%	17%
513 – Lakes and lagoons	**89%**	**61%**	11%	3%	**82%**	16%	3%
514 - Reservoirs	**100%**	**100%**	0%	1%	**99%**	0%	1%
523 – Sea and ocean	**80%**	**78%**	20%	25%	**69%**	6%	25%

The results of validation in the best scenario showed an average thematic accuracy of 73% (+/-8% SD) and geometric accuracy of 78% (+/-10% SD). Artificial classes got the lowest accuracy meanwhile forest classes obtained the highest accuracy: around 80% thematically and 85% geometrically. Thus LMAPP proves its potential as a solution to the automated mapping of afforestation and reforestation areas by identifying, at least, the forest type.

Nevertheless, our algorithm is designed in order to get a whole set of LC classes, including urban features, croplands, grassland, shrubs, sparsely vegetated land, wetland and water bodies' classes; beyond woodland and forests (Table 3.11, Figure 3.49). It is important to highlight that the result of algorithm depends on the requirements expected: in TIERRA 3 very high resolution LC map products are required so the resolution of the input data and the accuracy of the training samples areas should be optimal. However, for other LC products that requires other resolutions and scales, the algorithm LMAPP even with other input data (for example, Sentinel-2) could be a satisfactory solution.

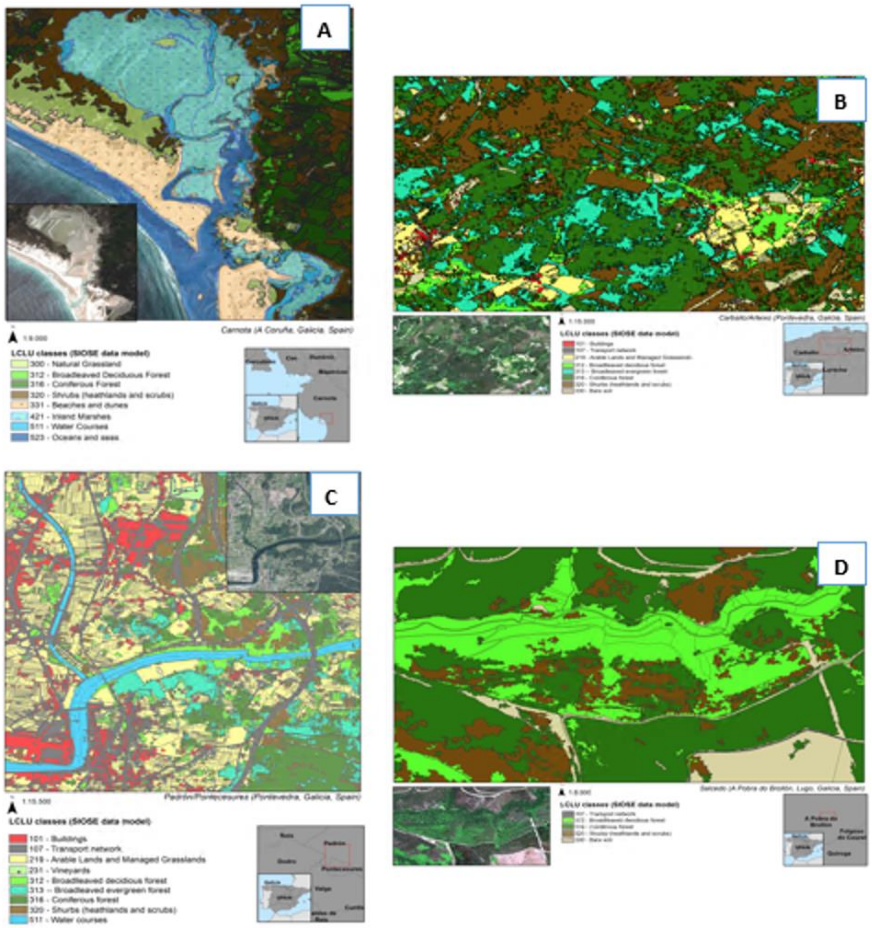


Fig. 3.49: Some examples of the results of the algorithm LMAPP: A) A Coastal marsh, B) An agroforestry area, C) A populated agroforestry area crossed by a river and connected by a road network, D) A woodland with areas of pruning. Examples B) and D) are attached to this report in shapefile format. Please notice that example B) is the result before generalization in order to show the possibility of detecting even littles areas of trees (despite commission and omission errors). In both cases only forest and shrub classes are presented to focus on the potential of the algorithm LMAPP to response the needs of the gap of LULUCF layer in matters of afforestation and reforestation areas.

<u>LMAPP algorithm: current progress</u>

Nowadays, COTESA keeps working on the algorithm LMAPP. On one hand, as a global LC mapping method, LMAPP is being testing in other regions of Spain. On this matter, the results suggest a paradigm shift by which a multi-strategical approach should be considered in order to improve the results - that means that different group of classes requires different methods to automate mapping. Thus, in the future LMAPP could be designed as a set of interconnected algorithms that work combined to produce very accurate LC map products. Additionally to that, and from the State-of-the-Art achieved during the development of the algorithm in the project TIERRA 3, our company keeps investigating and improving the method to adapt it to different sources, specifically to HR satellite imagery such as Sentinel-2. Next steps includes the validation of the results in order to determine the accuracy, the resolution and the output scale of LC map products derived of Sentinel-2 by LMAPP.

On the other hand, our company is applying the algorithm LMAPP in several projects related to forest cover and tree detection. It should be cited its applicability in electric corridor monitoring projects, in which our company generates vegetation alert maps by combining LMAPP with other in-house algorithms and using LiDAR data. Furthermore, recently the algorithm LMAPP has been applied in combination with change detection methods to evaluate damage of urban trees in Madrid after the snow-fall “Filomena” of January 2021 (Figure 3.50), achieving very accurate results in automatic detection of urban trees (Kappa=0.96).

Fig. 3.50: Detection of urban trees in Madrid by LMAPP before and after the snow-fall "Filomena" of January 2021. These results were the basis to performance damage and biomass loss assessment.

Additionally, COTESA is currently planning new tests of the algorithm LMAPP at tree species level in order to evaluate its current potential using HR and VHR imagery as input data.

LMAPP algorithm: future perspectives

To deal with the challenge of national or European scale and meet the needs of the gap identified in LULUCF of tree mapping in afforestation and reforestation areas, COTESA proposes to research the potential of Sentinel 2, despite the resolution, because of its high Near-Real Time latency and extensive coverage. In addition to that, it is needed to evaluate the potential of current LC map products (regional, national and European) as an accurate reference database to be used as training samples areas for the algorithm LMAPP. On this matter, a harmonized integration strategy of different databases (CLC, Forest HRL, CLMS Local products and regional or national database) should be researched, also the potential of future Land Copernicus products as CLC+. Furthermore, it is expected that algorithm LMAPP should be adapted for different biogeographical regions. Thus, new lines of research are foreseen in search of the optimization of the algorithm LMAPP and its applicability for national and European scales.

As mentioned before, the potential and improvement of LMAPP is currently being investigated in-house by adapting them to other different sources, other data models (including the species level based on multi-temporal series imagery) and testing in other regions of Spain. Related to that, COTESA intends in the long term to investigate the potential of hyperspectral data taken by STS spectroradiometer on-board an Unmanned Aerial Vehicle (UAV) and its applicability on deriving environmental information (Becerra *et al*., 2018) through PLSR techniques (i.e. spectral classes for each type of vegetation species). This methodology is key to understand the spectral behaviour and, hence, to retrieve environmental information by upscaling them to satellite-sensor-characteristics through spectral resampling. Thus, it is expected to achieve the trees species level with high accuracy.

Additionally to this line of research, COTESA has designed a super-resolution imaging algorithm supported by Deep Learning that keeps the spectral data, with promising results: the algorithm IRIX4 (Figure 3.51). By using algorithms based on convolutional neural networks CNNs and pre-processing satellite images, spatially super-resolved images are obtained from training with medium-resolution images. Maintaining the spectral character in the new 2.5-metre images allows more accurate classifications, segmentations and statistics to be obtained than with the original medium-resolution images. The use of GPUs in the training and execution stage offers an improvement in time and cost reduction for the achievement of a high added-value product. Thus, the future of LULUCF mapping could be addressed to the use of Sentinel -2 combined with super-resolution imaging techniques, getting VHR multispectral and multi-temporal data.

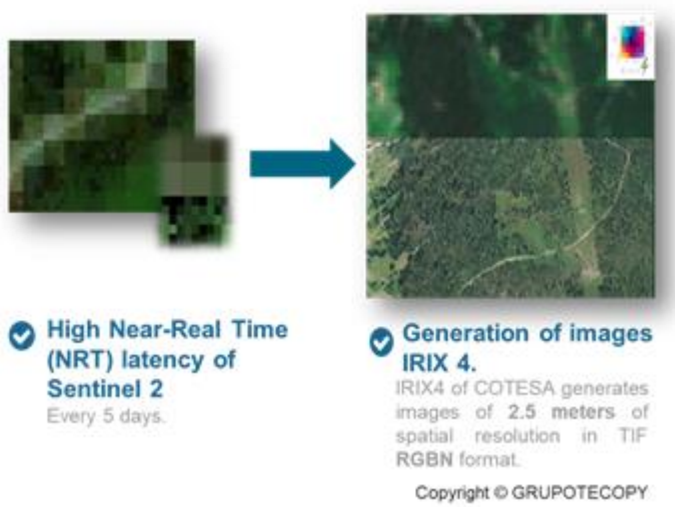


Fig. 3.51: Algorithm IRIX4 of COTESA.

Nowadays, the algorithm LMAPP is presented as a good candidate to improve the mapping of afforestation and reforestation areas by identifying, at least, the forest type and including a new level in the case of broadleaved forests, differentiating both deciduous and evergreen of broadleaved forests. Furthermore, LMAPP can improve the mapping of current LULUCF classes and support other gaps. The future lines of research will lead the challenge of reaching the tree species level and the applicability to national and European scales.

Key results

The algorithm LMAPP currently produce very accurate Land Cover maps, especially considering forest classes (coniferous forest, deciduous broadleaved forests and evergreen broadleaved forest). The future perspective points to the improvement of the classification getting the species level and the potential of Sentinel-2.

Innovative impact

Our algorithm can be applied to other gaps, not only related to Forest Layer and can support and improve the mapping of other LULUCF classes and current Land Copernicus products.

Related Copernicus domains: Land surface and vegetation.

Area of interest

The methodology was developed for several different zones of Galicia, on the northwest of Spain, covering more than 250 km^2 during the development phase (including areas tested with UAV, VHR and HR imagery). Additionally, other new areas were tested during the validation phase, covering over 200 km^2. Two samples are attached to this report: the results of the algorithm LMAPP detecting forest classes in two areas of Galicia based on two WV images and specific training samples (see Figure 21 B and D):

- The first sample (Figure 21 B) is located in Carballo (A Coruña), covering an area of 11,8 km^2 where plantations of coniferous (*Pinus* sp.) and evergreen broadleaved plantations (*Eucalyptus* sp.) are patched with deciduous broadleaved forests in an agroforestry area. It is presented the raw data before generalization in order to show the potential resolution and accuracy derived of the use of WV imagery.
- The second sample (Figure 21 D) is located in Salcedo (A Pobra do Bruñón, Lugo) covering an area of 3,7 km^2 where plantations of coniferous (*Pinus* sp.) and pruning areas are crossed by a riparian deciduous broadleaved forest. In this case the generalization has been performance.

Samples of other classes can be provided on demand.

Algorithm and data used

Pixel-based segmentations and RF classifications of multitemporal data from VHR and HR multispectral imagery, Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI).

Relevant literature

Ariza, A., Roa, O. J. y Serrato, P. K. y León Rincón, H. A. (2018). Uso de índices espectrales derivados de sensores remotos para la caracterización geomorfológica en zonas insulares del Caribe colombiano. Perspectiva Geográfica, 23(1), 105-122.

Becerra, J., Martin, M. P., Pacheco-Labrador, J., Gonzalez-Cascon, R., Melendo-Vega, J. R., & Angás, J. (2018). Chlorophyll Estimation in Mediterranean Quercus ilex tree canopies with hyperspectral vegetation indices at leaf and crown scales. YP Remote Sens, 7, 11.

Erb, K.-H., Kastner, T., Luyssaert, S., Houghton, R. A., Kuemmerle,T., Olofsson, P., and Haberl, H. (2013). Bias in the attribution of forest carbon sinks, Nature Climate Change, 3, 854–856.

Garcia-Oliva, F., Masera, O.R. (2004). Assessment and measurement issues related to soil carbon sequestration in land-use, land-use change, and forestry (LULUCF) projects under the Kyoto Protocol. Climatic Change 65, 347–364

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425-1432.

Xie, Y., Sha, Z., and Yu M. (2008). Remote sensing imagery in vegetation mapping: a review. Journal of plant ecology, 1: 9–23.

Success story #4 for LULUCF report in Spain. Remote sensing applied to mapping and estimation of current and past carbon stocks in seagrass meadows

Introduction

Seagrasses are carbon sinks, with a capacity comparable to that of terrestrial forests (*McLeod et al*. 2011). Their high sink capacity derives from high primary productivity that accumulates in the form of aboveground (leaves) and belowground (roots and rhizomes) biomass, as well as the efficiency of the seagrass canopy in trapping organic particles from the water column.

In recent years, many countries are working to incorporate seagrass into national greenhouse gas inventories and climate change mitigation strategies. However, making this intention a reality requires precise knowledge of the carbon stocks associated with these ecosystems and of the spatial scales relevant to the development of climate change mitigation strategies.

The carbon sequestered and stored by seagrasses varies across spatial and temporal scales, due to different structural and physiological characteristics of species, environmental conditions (e.g. irradiance, nutrient availability, water turbidity) and anthropogenic pressures. In a global context, where in situ assessment and monitoring of seagrass meadows and their carbon stocks requires an investment of resources that is often difficult to afford, remote sensing has shown its potential for identifying, mapping and monitoring changes in the surface area of coastal wetlands. It also has shown the potential for estimating vegetation variables closely related to the carbon sequestration capacity of these ecosystems (e.g. biomass, leaf area or photosynthetic activity) (Zhang *et al*., 1997; Zomer *et al*., 2009; Kuenzer *et al*., 2011).

The present case study demonstrates the use of remote sensing (i.e. satellite imagery) to map the extent of current and past seagrass meadows of *Zostera noltei* and to estimate the change in carbon stocks associated with their aboveground biomass in GHG emissions context. This general objective will be achieved through two specific objectives:

- To map the spatio-temporal changes in the area occupied by *Zostera noltei* intertidal seagrass meadows in the Bay of Santander in the period 1984-2015, through the analysis of images from Landsat (Landsat 4, 5, 7 and 8) and Sentinel -2 (A and B) satellites.

- To develop a predictive model of short-term carbon sequestration based on spectral images captured by Sentinel-2 and Landsat remote sensors and to analyse whether there is a correlation between vegetation indices and biophysical parameters calculated from satellite images and the C_{org} (organic carbon) content in the leaf biomass of *Zostera noltei* grasslands. The results obtained for Landsat may be applicable for retrospective studies of carbon sequestration, while the Sentinel results will be useful for studies of the near past and for follow-up work in the future.

Study area

The Bay of Santander (Spain) is one of the largest estuaries on the Cantabrian coast (2346 ha). It is classified as an estuary of complex typology in which tidal dynamics dominate (Galván *et al*., 2010). This estuary has a semi-diurnal tidal regime with an average tidal range of 2.8 m and a maximum of 4.9 m (Ports of the State). The maximum currents are of the order of 1.1 m/s and occur at the mouth, navigation channel, in front of the Puntal beach and in certain sectors of the Cubas estuary (Bidegain *et al*., 2015). The areas with the lowest velocities are located in the moorlands, harbour basins, Elechas and certain areas of the Cubas estuary (Figure 3.52).

The main freshwater supply comes from the river Miera. This river has an average annual flow of 8.2 m^3/s. The Bay of Santander also receives inflows from other smaller rivers located in its innermost area. It is an estuary in which intertidal areas predominate, covering around 67% of its total surface area. The sediment within the Bay is mainly sandy in the central and northern part and muddy in the southern and innermost areas (Bidegain *et al*., 2013). This estuary is occupied by extensive seagrass meadows, dominated by *Zostera noltei* in the intertidal zones and *Zostera marina* in the subtidal zones, an important part of which are included in the Special Area of Conservation (SAC) Puntal dunes and Miera estuary (Directive 92/43/EEC).

Fig 3.52: Detail of the study area in the Bay of Santander.

Cartography of Zostera noltei communities in cantabrian estuaries

Methodology

Collection of satellite images

For the retrospective study, images from the four Landsat programme satellites (Landsat 4, 5, 7 and 8) were used for the period 1984-2015. These images have a spatial resolution of 30 x 30 metres and an approximate revisit time frequency of 16 days. Each satellite uses different multispectral sensors (TM, ETM and OLI) that provide information in 6 ranges or bands of the spectrum between 450-2350 nm.

To capture the moment when the species is at its maximum stage of development, images obtained between the months of June and September were selected, with high luminosity, absence of clouds over the study area and a tide level as low as possible to maximise the surface of the emerged area. For the Bay of Santander 16 images were used. In the study period there are years in which suitable satellite images are not available during the months of June to September due to the presence of cloud cover, a high tide level, or defects in the sensor (e.g., banding of Landsat 7 images).

The satellite images used are specific surface reflectance products generated and distributed by the USGS (http://landsat.usgs.gov/CDR_LSR.php). Geometric and radiometric corrections are applied to the raw information provided by the remote sensors and, in the specific case of this product, the USGS applies the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS). LEDAPS is a specialised software that uses data from the MODIS (Moderate Resolution Imaging Spectroradiometer) product to perform the atmospheric corrections. Specifically, MODIS provides information on different atmospheric parameters. The effect of water vapour, ozone, aerosol optical thickness, etc., on the signal is estimated using the 6S (Second Simulation of a Satellite Signal in the Solar Spectrum) radiative transfer model. As the resulting product, the USGS provides information not only on surface reflectance, but also on cloud masks and cloud shadows among other additional data.

Pixel-based classification

A stepwise, or sequential, approach based on the establishment of thresholds or cut-off points and unsupervised classification techniques was used to classify the satellite images and identify the distribution area of *Z. noltei* (Calleja, 2017). The classification unit used was each pixel provided by the satellite images (Figure 3.53).

First, two types of categories were identified: pixels representing flooded areas and pixels representing emerged areas. To differentiate between the two classes, the NWI (Normalized Water Index) was used, which is able to detect the presence of water (Silió-Calzada *et al*., 2016). Positive values of the index indicate the presence of water and, therefore, pixels with such values were assigned to the class of flooded areas. Likewise, pixels with negative index values were assigned to the emerged zones class.

Subsequently, pixels in the emerged zones were classified into two categories: vegetated sediment and non-vegetated sediment. Specifically, pixels corresponding to non-vegetated sediment were identified using the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979). A threshold of 0.2 was established, so that pixels with an NDVI value lower than 0.2 were assigned to the non-vegetated sediment class and the rest to the vegetated sediment class (Barillé *et al*., 2010).

Next, pixels belonging to the vegetated sediment class (NDVI≥0.2) were classified into two classes: marsh vegetation and vegetation types other than marsh vegetation. For this, a principal component analysis (PCA) was applied with Landsat blue, red, green, near infrared (NIR) and mid-infrared 1 (SWIR1) bands and pixels with eigenvalues of the third principal component (PC3) greater than zero were selected. These pixels were then classified using the unsupervised k-means classification system and the Enhanced Vegetation Index 2 (EVI2) (Equation 1) (Ghosh *et al*., 2016) (Ghosh *et al*., 2016).

$$\mathbf{EVI2} = \frac{\mathbf{2.5 * (NIR - R)}}{\mathbf{(NIR + 2.4 * R + 1)}} \qquad \text{(Equation 1)}$$

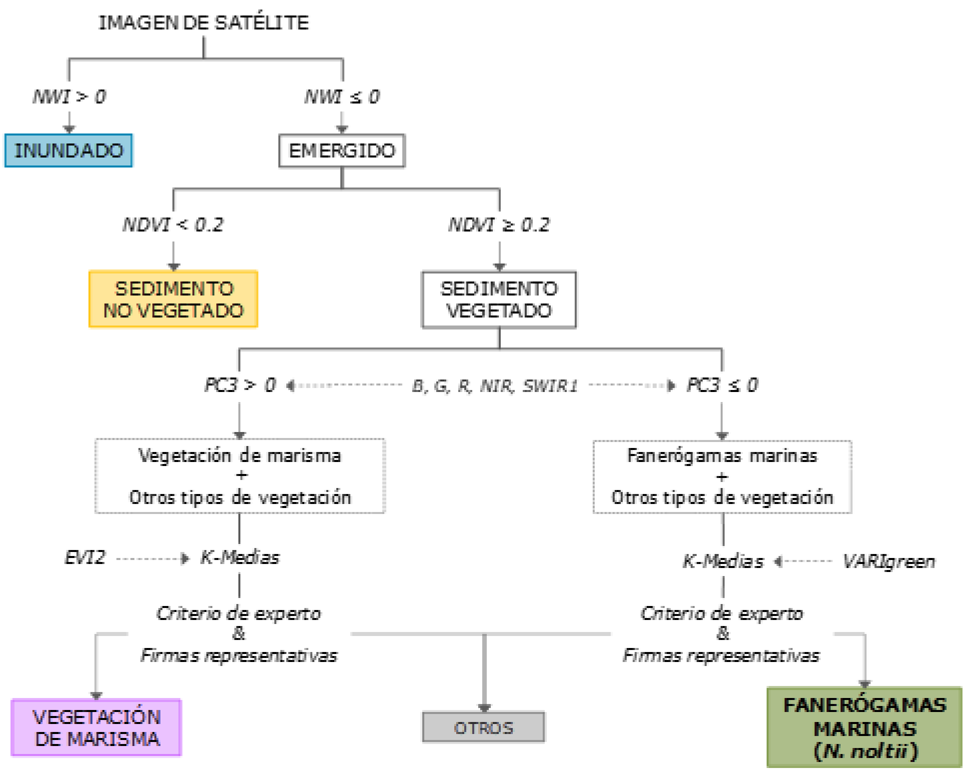
where 'NIR' is the near infrared band and 'R' is the red band (Landsat). Four groups were obtained as a result of the k-means analysis and the group(s) corresponding to marsh vegetation were identified by expert judgement based on the representative spectral signatures. The remaining groups were assigned to the class of other vegetation types.

Finally, pixels assigned to the class of vegetation types other than marsh vegetation were further classified into two classes: marine phanerogams and other vegetation types. Again, a k-means classification was performed, but with the Visible Atmospherically Resistant Index Green (VARIgreen) (Equation 2) (Gitelson *et al*., 2002).

$$\mathbf{VARIgreen} = \frac{\mathbf{G} - \mathbf{R}}{(\mathbf{G} + \mathbf{R} - \mathbf{B})} \qquad \text{(Equation 2)}$$

where 'G' is the green band, 'R' is the red band and 'B' is the blue band (Landsat).

Four groups were obtained as a result of the k-means analysis and the group or groups associated with the presence of marine phanerogams were identified on the basis of expert judgement and representative spectral signatures. The remaining groups were assigned to the class of other vegetation types, which includes macroalgae, diatoms, marsh vegetation not identified in the previous steps, etc.



NWI: Normalized Water Index; NDVI: Normalized Difference Vegetation Index; EVI2: Enhanced Vegetation Index 2; VARIgreen: Visible Atmospherically Resistant Index Green; PC3: principal component 3 from the principal component analysis; B, G, R, NIR, SWIR1: blue, green, red, near-infrared and mid-infrared bands of the Landsat satellite sensors.

Fig 3.53: Pixel-level classification system for range identification of Z. noltei based on a sequential methodology.

The spatial distribution models of *Z. noltei* constructed from the spectral information were compared with data on the real distribution obtained from field data in 2013. For this purpose, a Kappa analysis was carried out and a significance level of 95% was considered (Calleja, 2017).

Results

All Landsat images of the Bay of Santander for the period 1984-2018 were taken close to the high tide. Consequently, much of the estuary was submerged or at very high levels of humidity. This fact limited the application of the proposed methodology to identify the area occupied by *Z. noltei* to less than 20% of the estuary (Figures 3.54-3.56).

From the analysis of this 20% of the estuary, it can be seen that the distribution area of *Z. noltei* reaches its minimum in 2003 and its maximum between 2006 and 2015. In 2013, 2014 and 2015 the *Z. noltei* meadows were characterised by continuous meadows, which contrasts with those observed in previous years where the most characteristic distribution of the meadows is in aggregates or patches.

There are locations (i.e. marshes of El Conde and estuary of the Cubas) where the presence of meadows is intermittent, with years in which they practically disappear (e.g. 1987 or 1999) and locations where, on the contrary, throughout the whole historical series analysed, the presence and distribution of meadows is continuous (e.g. the Barquería). Depending on the year, the higher elevations of the Bay are colonised to a greater or lesser extent by *Z. noltei*; in this area there is great variability in the extent, morphology and type of arrangement of the meadows (i.e., in aggregates or continuous). In Elechas and Pontejos the most extensive and continuous meadows occur from 2006 onwards.

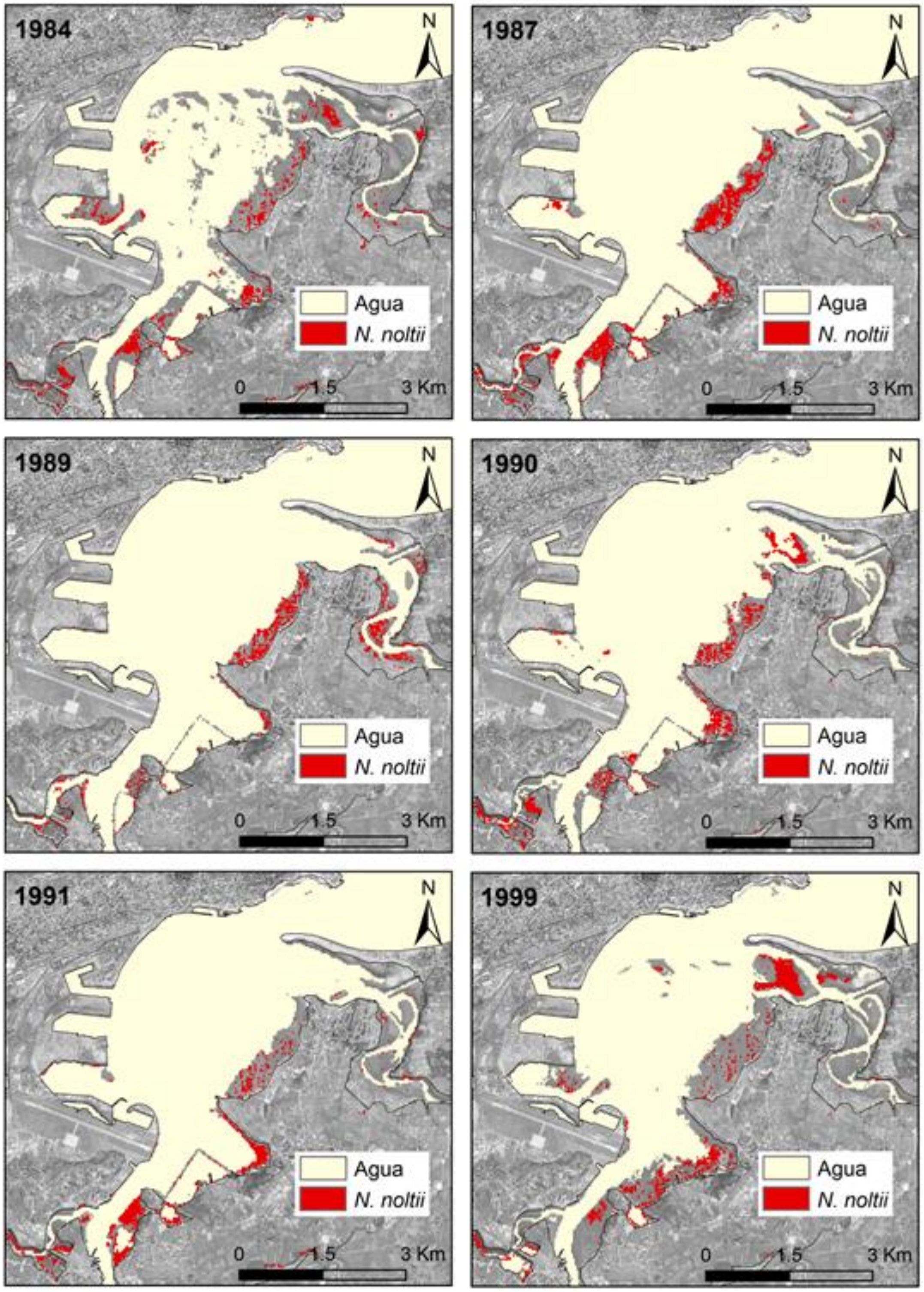


Fig 3.54: Distribution of *Z. noltei* in the Bay of Santander estuary in 1984, 1987, 1989, 1990, 1991 and 1999.

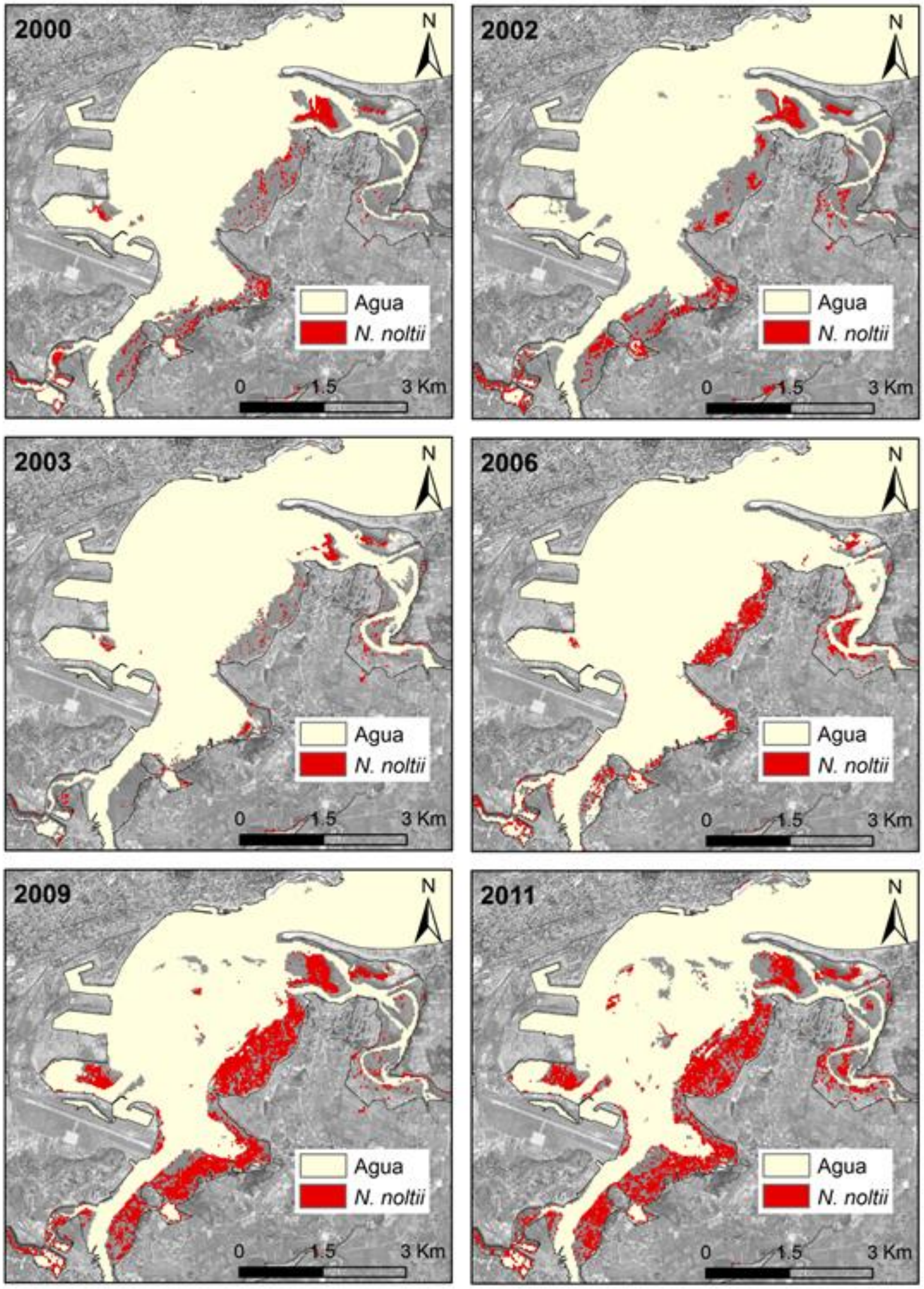


Fig 3.55: Distribution of *Z. noltei* in the estuary of the Bay of Santander in 2000, 2002, 2003, 2006, 2009 and 2011.

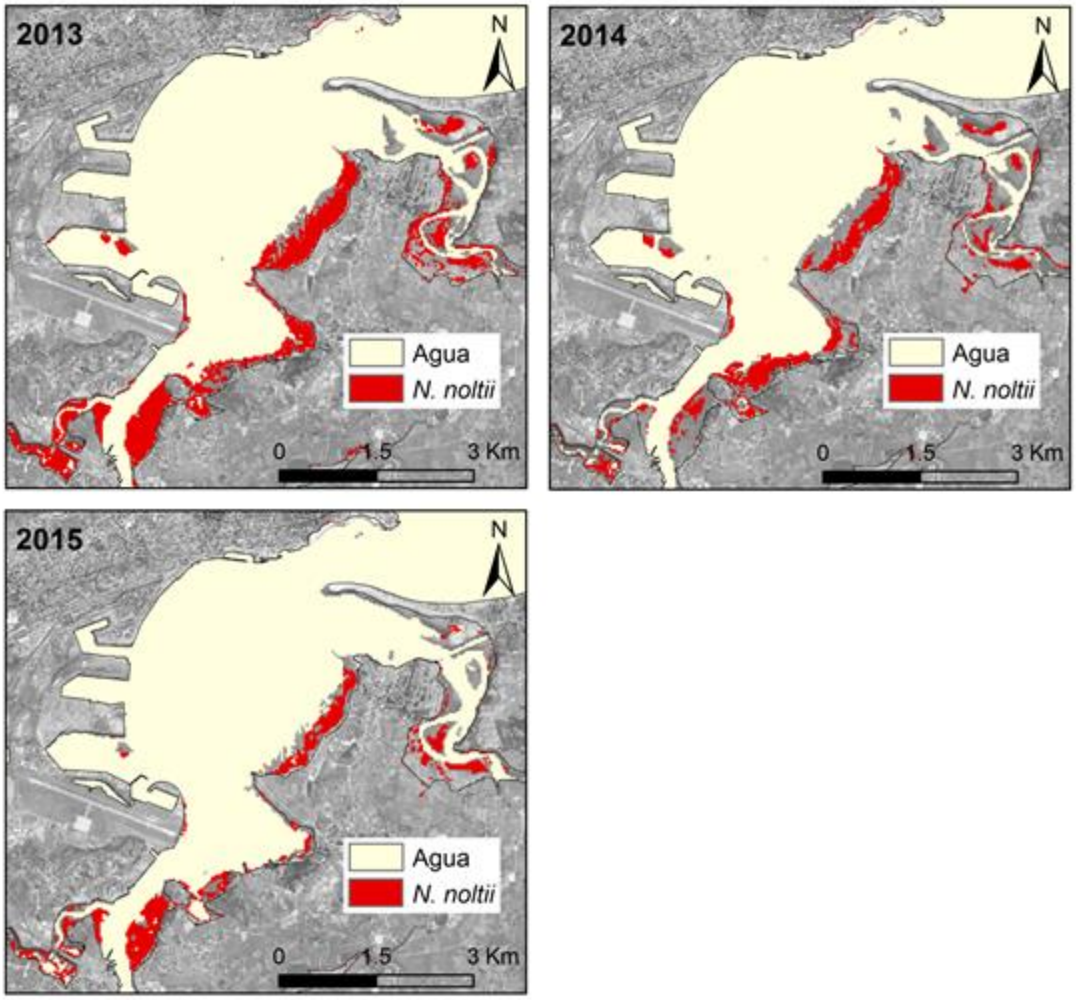


Fig 3.56: Distribution of *Z. noltei* in the Bay of Santander estuary in 2013, 2014 and 2015.

Estimation of the capacity of Zostera noltei cantabrian meadows for short-term carbon sequestration (deposits in live biomass)

Methodology

Carbon sequestration capacity of *Zostera noltei*

Six meadows were selected from the Bay of Santander (Cantabria). The short-term carbon sequestration capacity of *Zostera noltei* in biomass was measured as the C_{org} content in dry weight percentage and per unit area in aboveground biomass (or leaf biomass). Additionally, the C_{org} content in the living below-ground biomass (roots and rhizomes) of the meadows was quantified to assess the relationship between short-term and long-term carbon storage and in green algae, as previous studies show that it is difficult to discern by remote sensing between the presence of green algae and *Zostera noltei*.

At each meadow, a minimum of 3 to a maximum of 7 biomass samples were taken within an area of approximately 10 m diameter. Biomass samples were collected using biomass cores (ϕ 10-25 cm) down to approximately ~10-40 cm (the depth to which marine phanerogam roots reach) (Figure 3.57). Samples were processed followed the protocol described by Howard *et al*. (2014). Each sample was separated into different subsamples representative of the different components: aboveground or leaf biomass, live and dead belowground biomass of *Zostera spp*., green algae and detritus (Figure 3.58).

Fig. 3.57: Biomass sampling in the Bay of Santander. On the left, the biomass core used, spiking it to a depth of 20 cm. On the right, an example of the resulting sample in the mesh bag before being washed to remove sediment.

Fig. 3.58: Sample processing in the laboratory.

Differentiation between live and dead belowground biomass can be difficult in some cases, especially in small species such as *Zostera noltei*, and therefore the protocol of Howard *et al*. (2014) was followed. Aboveground biomass (live leaves) was immersed in 10% HCL for approximately 2 minutes to dissolve possible inorganic carbon content due to carbonate epiphytes. The subsamples were dried for a minimum of 72 h at 60ºC. After drying, they were placed in a desiccator for at least 2 hours until room temperature was reached and weighed (Dry Weight, DW).

The organic matter content of the subsamples was measured in all subsamples using the Loss on Ignition (LOI) technique, which consists of subjecting the sample to a temperature of 450ºC for two hours and estimating the organic matter content as the difference in weight of the sample before and after ignition (Gao *et al.*, 2003).

C_{org} was measured in 5 subsamples of live aboveground biomass (leaves), 5 subsamples of live belowground biomass (roots and rhizomes) and 5 subsamples of green algae from different meadows in the Bay of Santander using an elemental analyser (IHLab BIO laboratory of IHCantabria).

To estimate the concentration of C_{org} in the subsamples for which only organic matter was measured, the equation resulting from a linear regression analysis between organic matter concentration and organic carbon concentration were applied (Figure 3.59).

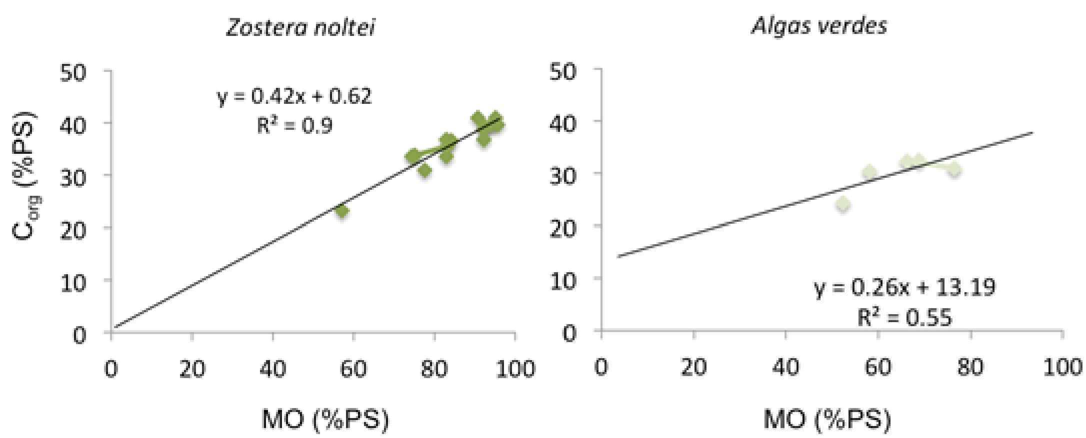


Fig 3.59: Ratio between organic matter (%PS) measured by LOI method and C_{org} concentration (%PS) measured with elemental analyser. (PS means dry weight).

The results obtained were used to estimate the biomass per unit area (g PS cm^{-2}), the concentration of organic matter and organic carbon both in dry weight percentage (%PS) and content per unit area (mg MO cm^{-2} , mg C_{org} PS cm^{-2}), of the different components considered: foliar aboveground biomass, live belowground biomass, dead belowground biomass and green algae.

For the estimation of the biomass of each component per unit area for each sample, the dry weight was divided by the sampled surface, calculated as $S_{sampled} = \Pi * r^2$, where "r" is the radius of the sediment core used in the sampling.

The organic matter and organic carbon content per unit area for each component was estimated by multiplying the concentration in %PS by the biomass per unit area. First, differences between meadow biomass components (aboveground, live belowground, dead belowground, green algae and detritus) were analysed in terms of dry weight per unit area and organic matter and organic carbon (in % dry weight and per unit area) for the whole meadow assemblage. Secondly, differences between meadows were analysed in terms of biomass (mg PS cm^{-2}) and organic matter and organic carbon content (%PS and mg PS m^{-2}) for the different meadow components. These differences were analysed by applying Analysis of Variance (ANOVA) and the Tukey-Kramer post-hoc test.

Predictive model for biological carbon sequestration

Landsat-8 surface reflectance images (applicable for retrospective analysis), and Sentinel-2 (applicable for the study of the near past, as well as for future follow-up studies) were selected for the study area with no clouds and corresponding to low tide periods. The Landsat-7 images captured for this area were faulty and were discarded.

Twenty-eight vegetation spectral indices were selected for Sentinel-2 (28 indices) and Landsat (18 indices) related to biophysical parameters associated with biological carbon storage (Table 3.12). In the case of Sentinel-2, these biophysical parameters were estimated directly with the "L2B biophysical processor" tool implemented in SNAP (Sentinel Application Platform):

- The leaf Area Index (LAI), which is used to predict photosynthetic primary production through the structure and function of the vegetation cover.

- Fraction of Vegetation Cover (FVC), which is used to estimate the amount of above-ground biomass and thus biological carbon stored through the structure and function of the vegetation cover.

- The Canopy Chlorophyll Content (Cab or CCC), which is an indicator of photosynthetic activity and thus of the primary productivity of the vegetation. The chlorophyll content of leaves is directly related to the ability of plants to assimilate CO_2.

- The Fraction of Absorbed Photosynthetically Active Radiation (fPAR) expresses the energy absorption capacity of the vegetation for photosynthesis and CO_2 capture, which depends on both vegetation structure and light conditions.

Table 3.12: Selection of vegetation spectral indices for Sentinel-2 and Landsat satellites.

INDEX	DESCRIPTION	FORMULA	FACTORS
SAVI	Soil Adjusted Vegetation Index	(1 + L) * (near_IR - red) / (near_IR + red + L)	L = 0.5
TSAVI	Transformed Soil Adjusted Vegetation Index	s * (near_IR - s * red - a) / (a * near_IR + red - a * s + X * (1 + s * s))	s = 0.5 a = 0.5 X = 0.08
MSAVI	Modified Soil Adjusted Vegetation Index	(1 + L) * (near_IR - red) / (near_IR + red + L)	L = 0.5
NDVI	Normalized Difference Vegetation Index	(near_IR - red) / (near_IR + red)	
MSAVI2	The second Modified Soil Adjusted Vegetation Index	(1/2) * (2 * (near_IR + 1) - sqrt((2 * near_IR + 1) * (2 * near_IR + 1) - 8 * (near_IR - * red)))	
DVI	Difference Vegetation Index	near_IR - red	
RVI	Ratio Vegetation Index	near_IR / red	
PVI	Perpendicular Vegetation Index	sin(a) * near_IR - cos(a) * red	a = 45
IPVI	Infrared Percentage Vegetation Index	near_IR / (near_IR + red)	
WDVI	Weighted Difference Vegetation Index	near_IR - g * red	g = 0.5
TNDVI	Transformed Normalized Difference Vegetation Index	sqrt(near_IR - red) / (near_IR + red) + 0.5)	
GNDVI	Green Normalized Difference Vegetation Index	(near_IR - green) / (near_IR + green)	

GEMI	Global Environmental Monitoring Index	eta * (1 - 0.25 * eta) - (red - 0.125) / (1 - red) eta = (2 * (near_IR * near_IR - red * red) + 1.5 * near_IR + 0.5 * red) / (near_IR + red + 0.5)	
ARVI	Atmospherically Resistant Vegetation Index	(near_IR - rb) / (near_IR + rb) rb = red - gamma * (blue - red)	blue: B2 gamma = 1
NDI45	Normalized Difference Index	(near_IR - red) / (near_IR + red)	
MTCI	Meris Terrestrial Chlorophyll Index	(near_IR - red2) / (red2 - red1)	
MCARI	Modified Chlorophyll Absorption Ratio Index	((red2 - red1) - 0.2 * (red2 - green)) * (red2 / red1)	
REIP	Red-Edge Inflection Point index	700 + 40 * ((red1 + near_IR)/2) - red2) / (red3 - red2))	
S2REP	The Sentinel-2 Red-Edge Position index	705 + 35 * ((red1 + near_IR)/2) - red2) / (red3 - red2))	
IRECI	Inverted Red-Edge Chlorophyll Index	(near_IR - red1) / (red2 / red3)	
PSSRA	Pigment Specific Simple Ratio (chlorophyll) index	near_IR / red	
SIPI	Structure Intensitive Pigment Index	(800nm - 445nm) / (800nm + 680nm)	
EVI	Enhanced Vegetation Index	2.5 * (near_IR - red) / ((near_IR + 6*red - 7.5*blue) + 1)	
EVI2	Enhanced Vegetation Index 2	2.4 * (near_IR - red) / (near_IR + red + 1)	
CHL-RED-EDGE	Chlorophyll Red Edge	([760:800] / [690:720]) ^ (-1)	

ARI1	Atmospherically Resistant Vegetation Index	(1 / 550nm) – (1 / 700nm)
CI green	Chlorophyll Index Green	(near_IR / green) -1
CI red edge	Chlorophyll Index Red Edge	(near_IR / red) -1

The selection criteria for the indices were: 1) To show a significantly high correlation with the C_{org} content in leaf biomass (both as dry weight and per unit area) for the meadows examined, but no significant correlation with the C_{org} content in green algae; and 2) to show a lineal correlation between the vegetation index and the C_{org} content.

For this purpose, the correlation between each of the vegetation indices calculated at their highest resolution (10*10 m for Sentinel-2 and 30*30 m for Landsat) and the C_{org} content, both as a percentage of dry weight (%PS) and per unit area (mg C_{org} cm^{-2}), was assessed for the above-ground biomass of *Zostera spp.* and for the green algae of the meadows examined. These correlations (a total of 244) were assessed by Pearson correlation analysis using the Hmisc package of the R Studio software (Version 1.1.463).

Once the indices were chosen, the predictive model for each of them was established by fitting the index vs. carbon sequestration variable relationship to a linear equation by means of a linear regression analysis using the JMP 10.0.0 program.

Images obtained during the summer (June-September) and the resulting mapping from Chapter 3 were used and the predictive models developed with the chosen spectral indices for Landsat were applied to each of the image pixels (30 * 30 m) to assess the change in C_{org} sequestration in leaf biomass per unit area (mg C_{org} cm^{-2}) between 1985 and 2015 (retrospective analysis).

The predictive models with the chosen spectral indices for Sentinel-2 were applied to the relative assessment of C_{org} sequestered in leaf biomass. Satellite images obtained with Sentinel-2 in July 2018 and the mapping of *Zostera noltei* obtained for the year 2018 (Chapter 3) were used. Chosen spectral indices were calculated for Sentinel-2 for pixels of 10 * 10 m resolution. Meadow C_{org} content per unit area (mg C_{org} cm^{-2}) was estimated for each pixel by applying the two predictive models developed.

<u>Result</u>

The indices selected to establish the predictive models were GNDVI and ARVI for Landsat imagery and RVI and ARVI for Sentinel-2 imagery (Table 3.13).

The significant relationship between these indices and the seagrass aboveground biomass C_{org} sequestration variables (Figures 3.60 and 3.61) was analysed by means of a linear regression analysis whose results are shown in Table 14. According to this analysis, the vegetation indices obtained by both satellites explain better the variance between meadows in terms of carbon content per dry weight (C_{org} %PS) (R^2= 0.35 -0.85) than per unit area (mg C_{org} cm^{-2}; R^2= 0.21-0.31). However, the C_{org} content per unit area better reflects the C_{org} sequestered in the meadow, as it does not only consider the C_{org} concentration in the leaves but also the leaf biomass of the meadow. Therefore, the predictive models chosen to carry out the retrospective and spatial analysis of C_{org} sequestration (next section) correspond to the linear equations obtained from the linear regression analysis between vegetation indices and the C_{org} content per unit area (mg C_{org} cm^{-2}).

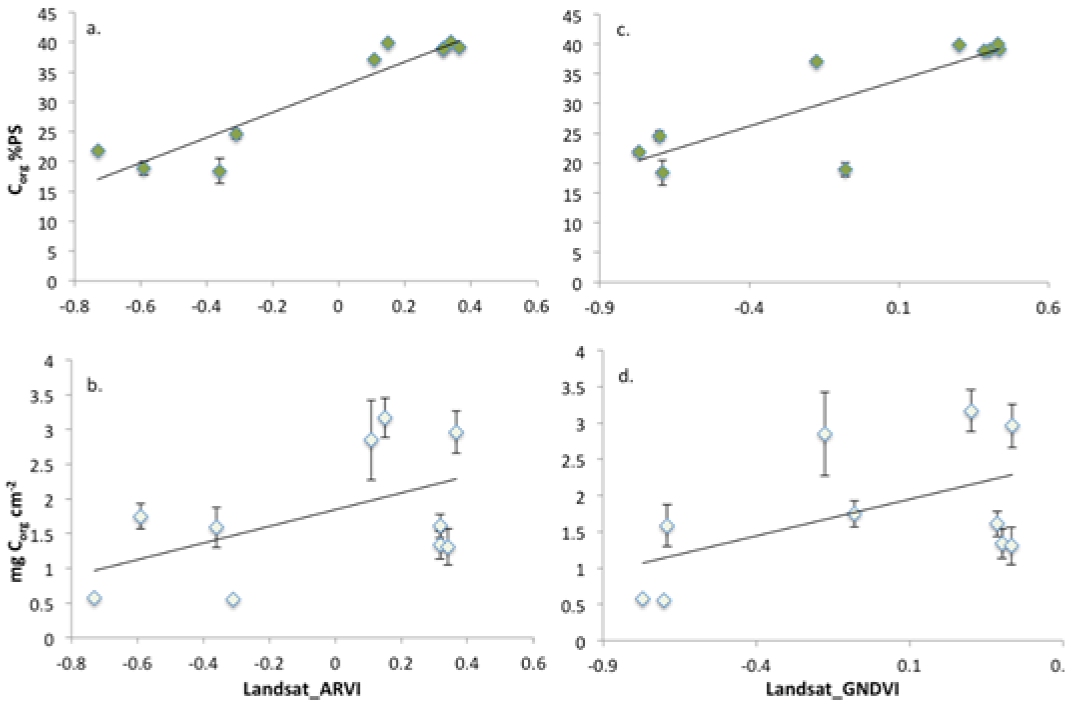


Fig. 3.60: Relationship between vegetation indices chosen for Landsat and C_{org} content in percent dry weight (a,c) and per unit area (b,d) (mean ± standard error). The lines represent the fit to the linear equations obtained from the regression analyses (Table 3.13).

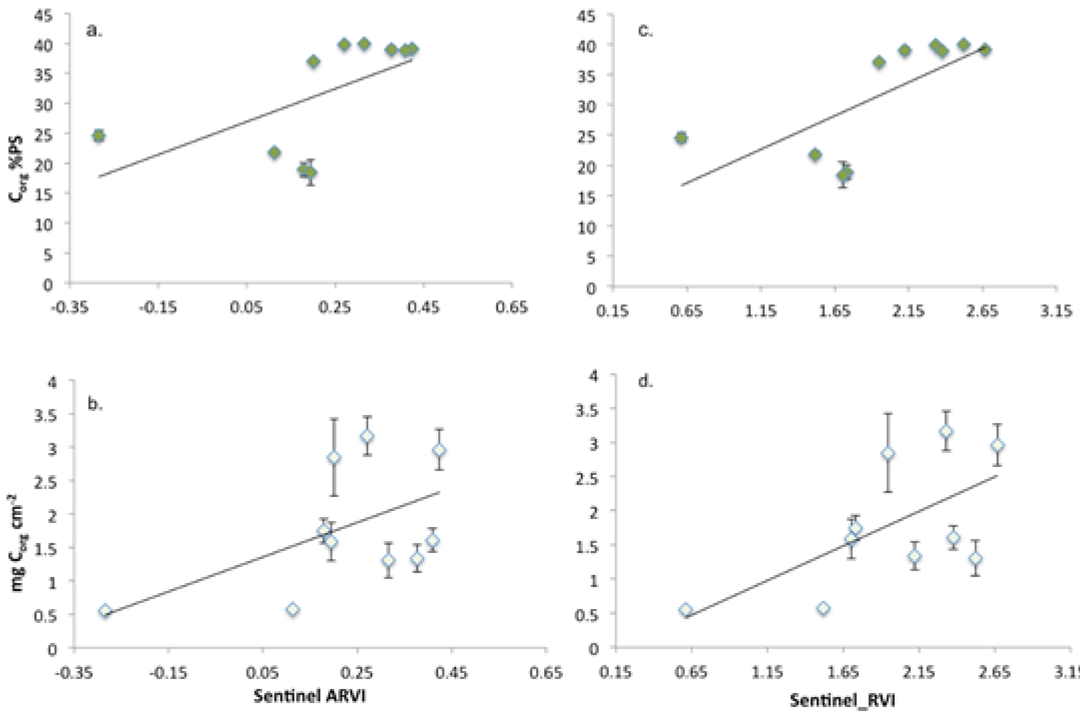


Figure 3.61: Relationship between the vegetation indices chosen for Sentinel-2 and C_{org} content in dry weight percentage (a,c) and per unit area (b,d) (mean ± standard error). The lines represent the fit to the linear equations obtained from the regression analyses (Table 3.13).

Table 3.13: Results of the linear regression analyses between the spectral indices chosen for each satellite and the carbon sequestration variables (C_{org} %PS, mg C_{org} cm^{-2}). The equations resulting from the linear fit constitute the predictive models of C_{org} sequestration.

Satellite	Regression analysis	Model	R^2	p
Landsat	GNDVI vs. C_{org} %PS	C_{org} %PS = 32.41 + 15.51*GNDVI	0.68	<0.0001
	GNDVI vs. mg C_{org} cm^{-2}	mg C_{org} cm^{-2} = 1.85 + 1.011*GNDVI	0.21	<0.001
	ARVI vs. C_{org} %PS	C_{org} %PS = 32.46 + 21.14*ARV	0.85	<0.0001
	ARVI vs. mg C_{org} cm^{-2}	mg C_{org} cm^{-2} = 1.84 + 1.20*ARVI	0.2	0.001

Sentinel-2	RVI vs. C_{org} %PS	C_{org} %PS = 9.90+ 11.09*RVI	0.5	<0.0001
	RVI vs. mg C_{org} cm^{-2}	mg C_{org} cm^{-2} = -0.19 + 1.01*RVI	0.31	<0.0001
	ARVI vs. C_{org} %PS	C_{org} %PS = 25.56+ 27.56*ARVI	0.35	<0.0001
	ARVI vs. mg C_{org} cm^{-2}	mg C_{org} cm^{-2} = 1.23 + 2.59*ARVI	0.23	0.0005

Predictive models developed from Landsat indices (GNDVI and ARVI) were applied to estimate C_{org} sequestered in *Zostera noltei* meadows from past satellite images (retrospective analysis). Predictive models developed from Sentinel-2 indices (RVI and ARVI) were applied to the relative analysis of C_{org} sequestered in meadows in the present (spatial analysis). However, given the low predictive power of the models (R^2=0.21-031), the results of the application of the latter should be interpreted with caution (or as a first approximation).

Retrospective analysis of C_{org} sequestration

Figures 3.62 and 3.63 show the estimated change in C_{org} content in aboveground biomass per unit area in *Zostera noltei* meadows for the years 1984-2015, applying the predictive models developed for the spectral indices GNDVI and ARVI calculated from Landsat at 30*30 m resolution.

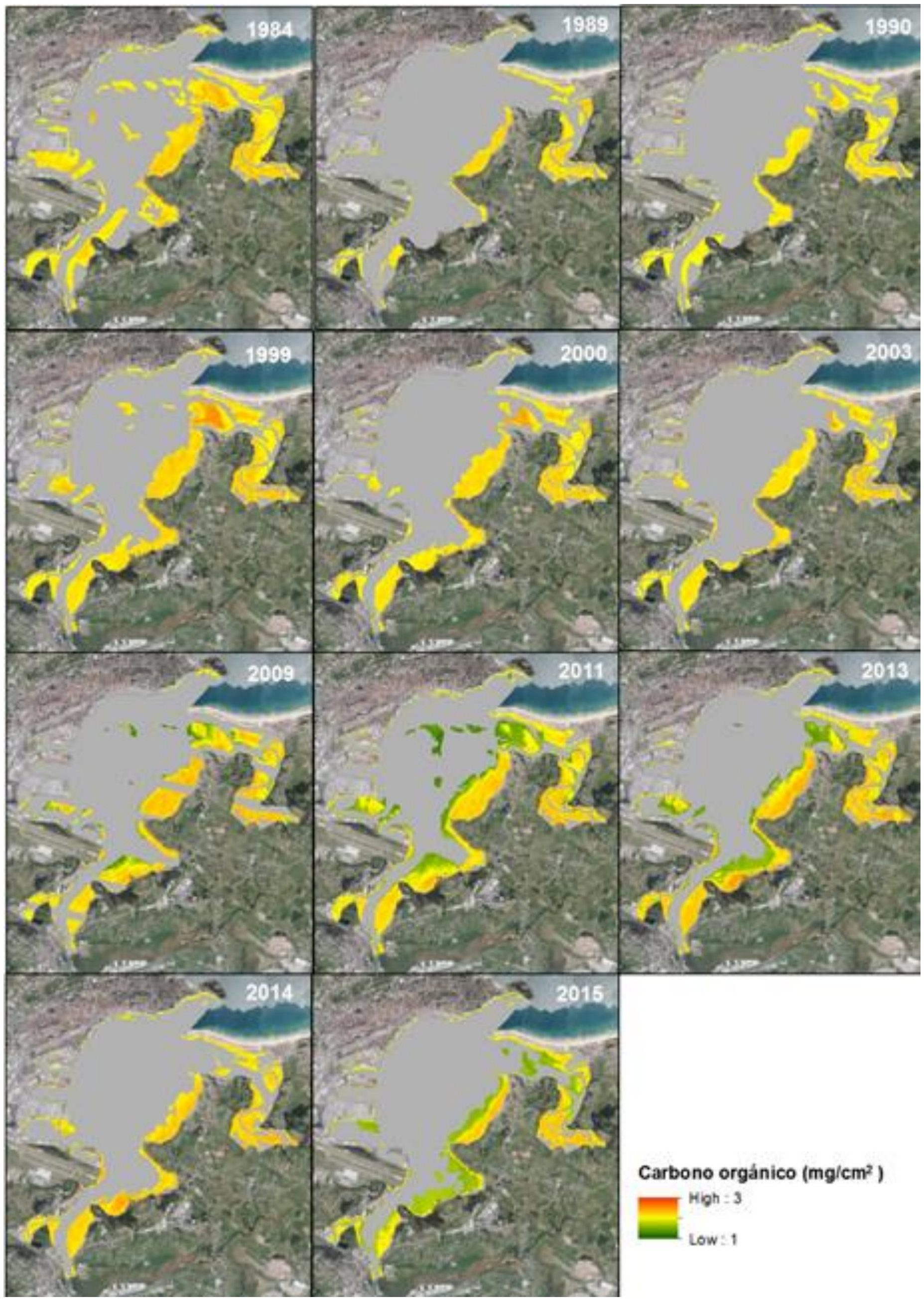


Fig 3.62: Evolution of the C_{org} content in aboveground biomass per unit area of *Zostera noltei* meadow in the Bay of Santander estimated on the basis of the GNDVI index obtained from satellite images captured with Landsat and applying the predictive model $mg_{Corg}\ cm^{-2} = 1.85 + 1.011*GNDVI$.

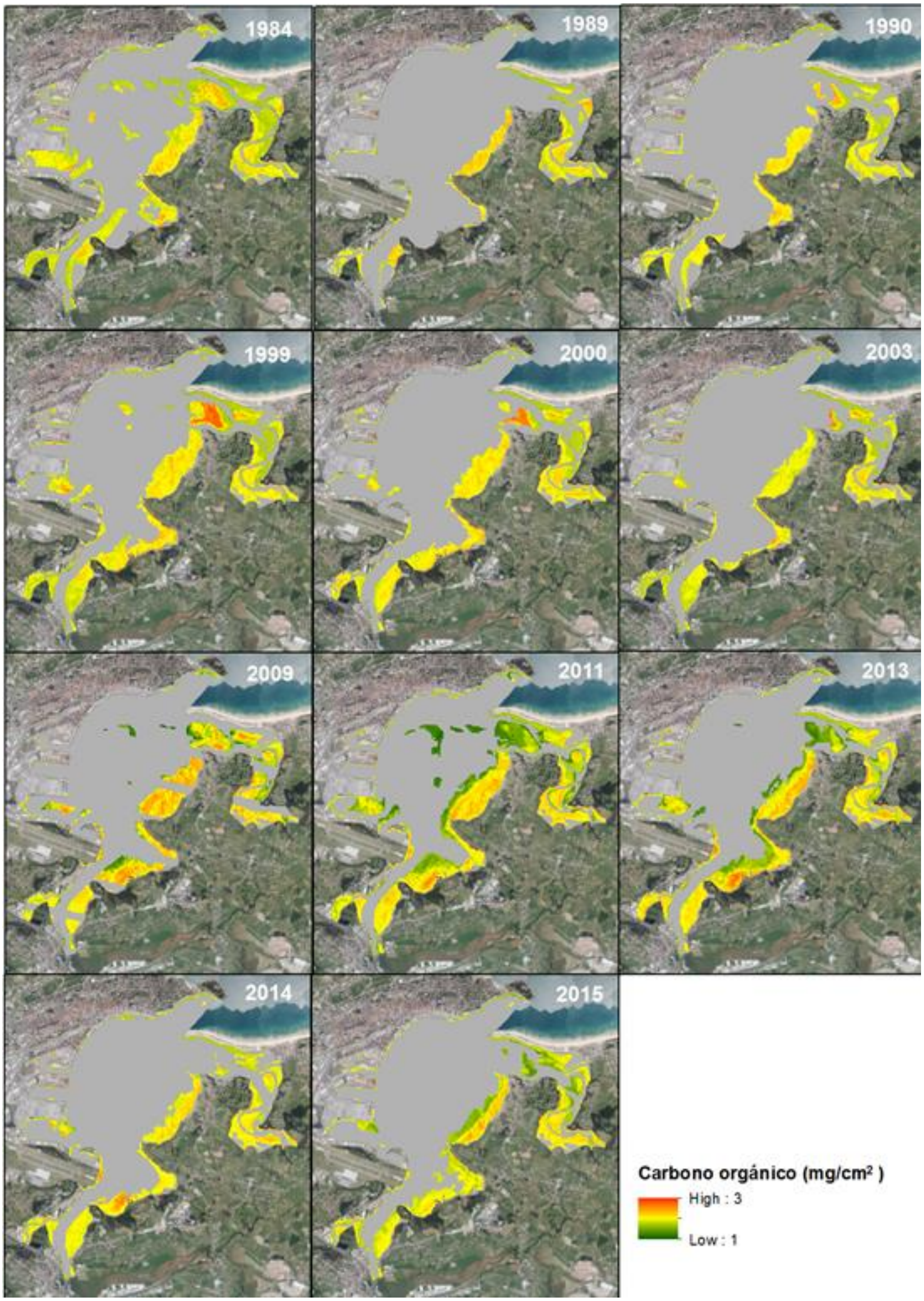


Fig. 3.63: Evolution in the C_{org} content in aboveground biomass per unit area of *Zostera noltei* meadow in the Bay of Santander estimated on the basis of the ARVI index obtained from satellite images captured with Landsat and applying the predictive model mg C_{org} cm-2 = 1.84 + 1.20*ARVI.

https://www.copernicus-user-uptake.eu/

Current analysis of C_{org} sequestration

Figures 3.64 and 3.65 show the spatial distribution of C_{org} content in aboveground biomass per unit area in *Zostera noltei* meadows in the year 2018. These values are estimated with the predictive models developed for RVI and ARVI spectral indices calculated with Sentinel-2 at a resolution of 10*10 m.

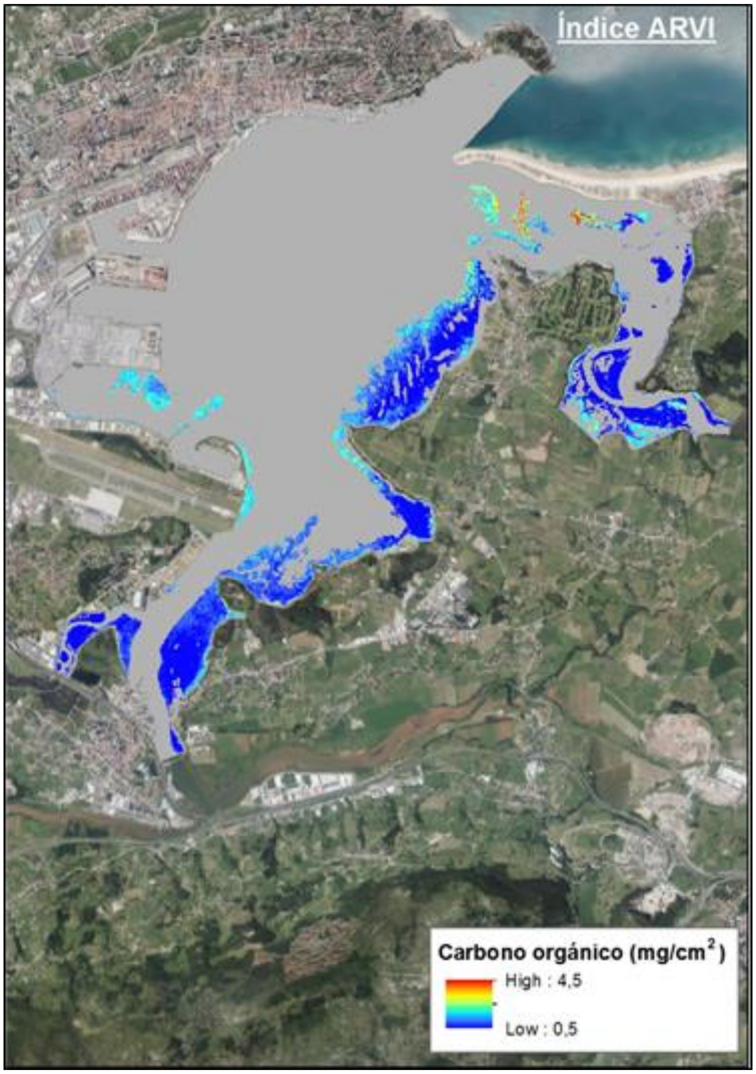


Fig. 3.64: C_{org} content in above-ground biomass per unit area of *Zostera noltei* meadow in the Bay of Santander estimated on the basis of the ARVI index obtained from satellite images captured with Sentinel-2 applying the model mg C_{org} cm^{-2} = 1.23 + 2.59*ARVI.

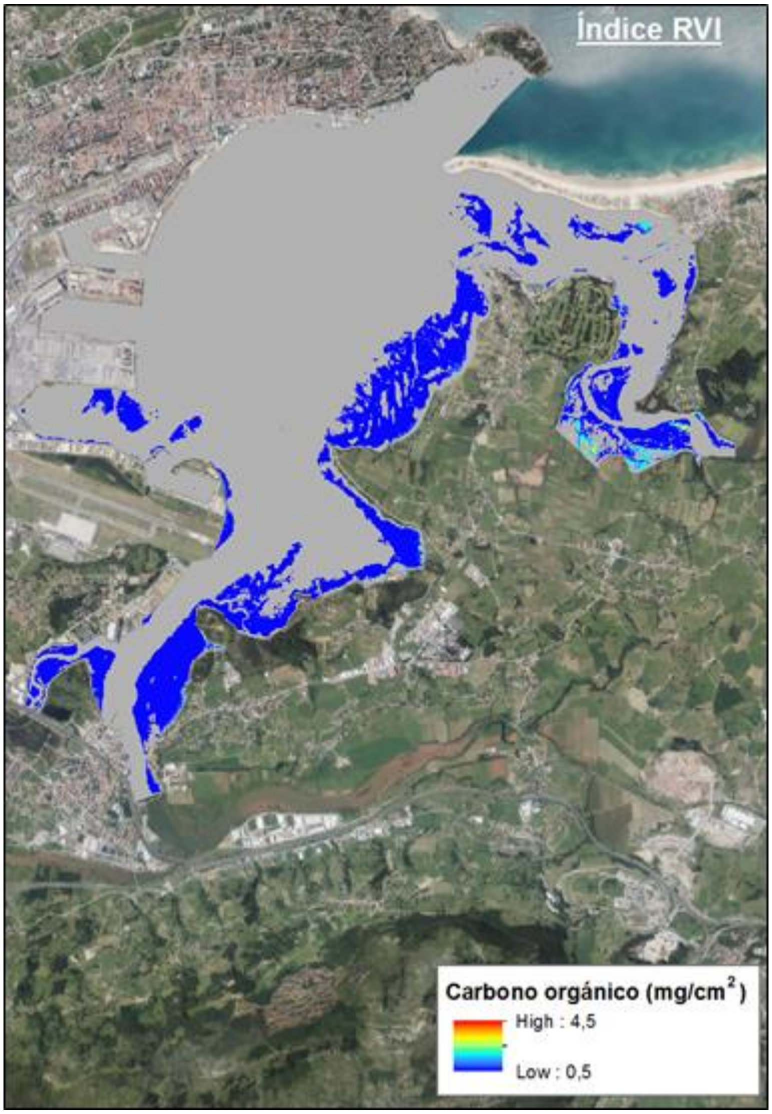


Fig. 3.65: C_{org} content in aboveground biomass per unit area of *Zostera noltei* meadow in the Bay of Santander estimated on the basis of the RVI index obtained from satellite images captured with Sentinel-2 applying the model mg C_{org} cm^{-2} = -0.19 + 1.01*RVI.

Conclusions

The methodology developed using satellite images has demonstrated its capacity to characterise the distribution area of *Z. noltei* in an objective, homogeneous and systematic way. These characteristics constitute an added value for the cartographic works derived from the application of this type of approach.

This approach makes it possible to carry out studies that cannot be carried out with fieldwork, such as retrospective studies and the permanent updating of the cartographies through the satellite images that are received periodically. However, estuaries are environments exposed to tidal action and with high levels of humidity that interfere with the spectral signal of the different elements of the estuary. This effect and the high variability in the environmental conditions of these systems make the interpretation of satellite images very difficult.

Remote sensing techniques have shown great potential to detect, identify, map and monitor the conditions and changes in the extent of other natural carbon sinks such as terrestrial forests, or even mangroves and wetlands. In addition, the most advanced remote sensing techniques allow estimation of vegetation variables closely related to their short-term carbon sequestration capacity such as biomass, leaf area or photosynthetic activity through the calculation of different spectral indices (e.g. NDVI). The subtidal condition of most seagrass meadows limits the application of such techniques. However, in the case of intertidal species, such as *Zostera noltei*, it has been shown that it is possible to map their extent in the emerged intertidal and retrospectively analyse changes in their extent (Calleja *et al*., 2017). However, the possibility of applying remote sensing techniques to the assessment of C_{org} associated with seagrass meadows has been little explored so far. In this sense, this experience represents a first important step in the application of these techniques to the assessment of temporal and spatial changes in short-term C_{org} sinks (in live leaf biomass) associated with seagrass meadows.

References

Bidegain, G., Bárcena, J. F., García, A., & Juanes, J. A. (2013). LARVAHS: Predicting clam larval dispersal and recruitment using habitat suitability-based particle tracking model. Ecological Modelling, 268, 78–92.

Bidegain, G., Bárcena, J. F., García, A., & Juanes, J. A. (2015). Predicting coexistence and predominance patterns between the introduced Manila clam (*Ruditapes philippinarum*) and the European native clam (*Ruditapes decussatus*). Estuarine, Coastal and Shelf Science, 152(162-172).

Calleja, F., Galván, C., Silió-Calzada, A., Juanes, J.A., Ondiviela, B.2017. Long-term analysis of *Zostera noltei*: A retrospective approach for understanding seagrasses' dynamics. Marine Environmental Research,130, pp. 93-105.

Galván, C., Juanes, J. A., & Puente, A. (2010). Ecological classification of European transitional waters in the North-East Atlantic eco-region. Estuarine, Coastal and Shelf Science, 87, 442–450.

Gao, X.; Huete, A.R. y Didan, K. (2003): Multisensor comparisons and validation of MODIS vegetation indices at the semiarid Jornada Experimental Range. IEEE Transaction on Geoscience and Remote Sensing, 41 (10): 2368-2381.

Howard, J., Hoyt, S., Isensee, K., Pidgeon, E., & Telszewski, M. (2014). Coastal Blue Carbon. Conservation International (Vol. 36).

Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T.V., & Dech, S. (2011). Remote Sensing of Mangrove Ecosystems : A Review.

McLeod, E., Chmura, G. L., Bouillon, S., Salm, R., Björk, M., Duarte, C. M., Lovelock, C. E., Schlesinger, W. H., & Silliman, B.R. (2011). A blueprint for blue carbon: Toward an improved understanding of the role of vegetated coastal habitats in sequestering CO_2. Frontiers in Ecology and the Environment, 9(10), 552–560.

Silió-Calzada, A., Barquín, J., Huszar, V. L. M., Mazzeo, N., Méndez, F., & Álvarez-Martínez, J. M. (2016). Climatic and anthropogenic effects on the long-term dynamics of a floodplain shallow lake in the Pantanal wetland.

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

Zhang, M., Ustin, S.L., Rejmankova, E., & Sanderson, E.W. (2019). Monitoring Pacific Coast Salt Marshes Using Remote Sensing. Ecological Applications, 7(3), 1039–1053.

Zomer, R.J., Trabucco, A., & Ustin, S.L. (2009). Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. Journal of Environmental Management, 90(7), 2170–2177. Retrieved from http://dx.doi.org/10.1016/j.jenvman.2007.06.028.

Success story #5 for LULUCF report in Spain. Modeling Soil Organic Carbon from the outputs of REMOTE

Keywords: soil conservation, vertical distribution of carbon, uncertainties map, ensemble of models, machine learning.

Application field: estimation of spatial soil organic carbon.

Abstract

Quantification and monitoring of soil organic carbon is needed to assess soil management practices, adapt policies, and evaluate environmental impacts. However, due to SOC spatial variability, soil surveys are a very challenging task with respect to the high costs to acquire such data, operational complexity, and updating. Remote sensing indices, in combination with machine learning approaches, have revealed an important improvement of soil carbon spatial distribution even when there exist limited soil samples. Environmental data-driven covariates intimately related to soil forming factors (e.g. biota, climate, parent material…) can be derived by spectral indices. Thus, we are working on the development of SOC spatial prediction modeling based on digital soil mapping techniques using remote sensing indices among predictors. We applied this modeling to generate a cost-effective, highresolution map of SOC distribution, and its associated spatially explicit uncertainty, for six depth layers (0-5, 5-15, 15-30, 30-60, 60-100 and 100-200 cm). This method aim to improve the benchmark SOC estimates to support the National Inventory Report of GHG emissions.

Methodology used / workflow

We estimated the spatial distribution of soil organic carbon (SOC) density, and its associated spatially explicit uncertainty, using state-of-the-art and emerging technologies for soil mapping. This study was performed within the framework of GlobalSoilMap (GSM) with data mining techniques based on the relevant soil forming environmental factors to generate a continuous map form sample points (Figure 3.66).

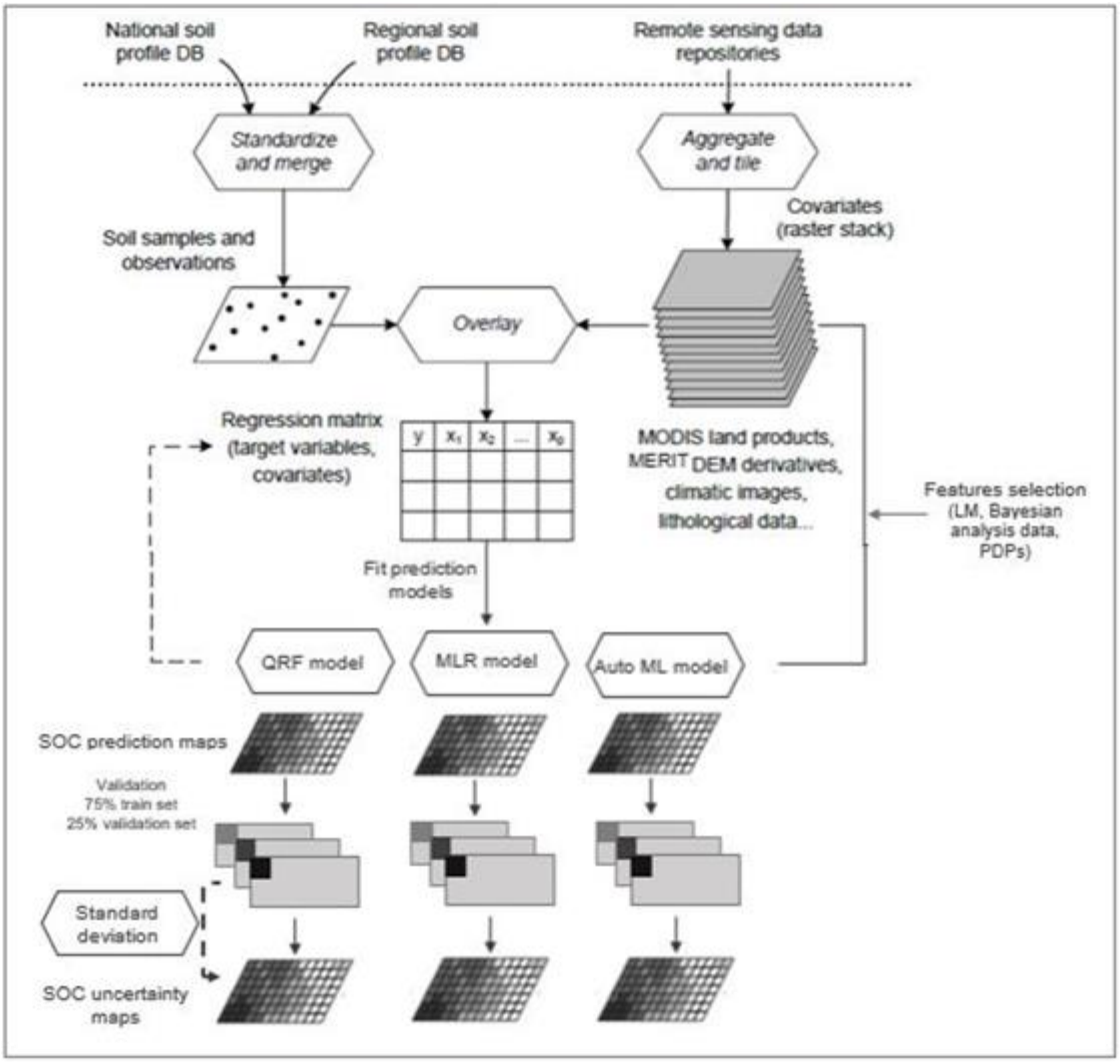


Fig. 3.66: Methodological scheme adapted from The World Soil Information Service (WoSIS).

The sample point database used in the current case study (Andalusian region) had a total of 1551 legacy soil profiles gathered by “Junta de Andalucía (CMAyOT)”. Prior to modeling, the harmonization and screen of soil data were necessary.

To predict the soil carbon as a function of the soil forming environment, we used the SCORPAN spatial inference model (Florinsky, 2012). According to the model, the environmental data-driven factors considered as SOC predictive covariates were related to terrain morphometric parameters, climate, land use and satellite-derived ecosystem functional attributes (EFA). We assessed the topographical relief through geomorphometry, and feature extraction derived from the Geomorpho90m global dataset, at 90 m resolution under the WGS84 geodetic datum. The fully standardized 26 geomorphometric variables of the dataset, derived from the MERIT-Digital Elevation Model (DEM), consist of layers that describe the (i) rate of change across the elevation gradient, using first and second derivatives, (ii) ruggedness, and (iii) geomorphological forms (Amatulli *et al*. 2020). The 1971- 2000 time series of climatic variables (mean monthly temperature and total monthly precipitation) at 100 m pixel resolution come from the REDIAM-Junta de Andalucía. The land use layer (updated to 2007) was derived from the same information net. In the categorical variables, the number of soil samples across all the categories were balanced, eliminating those with less than 100 data. The rest of the classes were transformed into dummies variables where each category in the map becomes an independent binomial predictor variable (e.g., only 0 and 1 values) as is explained in (Yigini *et al*., 2018).

The integration of remote sensing information as surrogate of different soil forming factors improved the SOC prediction capacity (Fathololoumi *et al*., 2020; Peng *et al*., 2020; Schillaci *et al*., 2017). Thus, we selected five satellite-products for the 2001-2019 period related to different dimensions of soil ecosystem functioning such as:

- Carbon cycle dynamics: Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI)

- Sensible heat dynamics: Land Surface Temperature (LST)

- Radiative balance: Albedo

- Water cycle: Precipitation and Normalized Difference Water Index (NDWI).

The sources and resolution of these satellite-products are described in Table 3.14.

Table 3.14: Description of satellite-derived ecosystem functional attributes.

Dimensions of ecosystem functioning	Spectral indices	Satellite-Product	Spatial resolution
Carbon cycle	**NDVI** y **EVI**: Normalized Difference Vegetation Index y Enhanced Vegetation Index	MOD13Q1	250 m
Water cycle	**Pre**: Precipitation **NDWI**: Normalized Difference Water Index	CHIRPS MCD43A4	1km 500 m
Radiative balance	**Alb**: Albedo	MCD43B3	500 m
Sensible heat	**LST**: Land Surface Temperature	MOD11A2	1 km
Ecosystem Functional attributes			
Annual total amount: mean, maximum, minimum	**Seasonality**: *standard deviation, coefficient of variation, range, relative range*	**Phenology**: *sine and cosine of the dates of maximum and minimum*	

Google Earth Engine (Gorelick *et al*., 2017) was used to derive the inter-annual mean and monthly mean of the following eight summary metrics of their seasonal dynamics: annual mean (surrogate of annual total amount), annual maximum and minimum (indicators of the annual extremes), seasonal standard deviation (descriptor of seasonality), and sine and cosine of the dates of maximum and minimum (indicators of phenology) (Alcaraz-Segura *et al*., 20017) The complete remote sensing dataset included 172 ecosystem functioning indices as candidate predictors.

We prepared the covariates for predictive models by the standardization of coordinate reference system (WGS84, EPSG: 4326) and spatial resolution (250-m grid). In order to reduce the geometric distortions in this process, we generated a covariate matrix based on the pixel centers of the MODIS resolution (i.e., 250 m pixel resolution). Each column in the matrix contains the values of a covariate that correspond to the same location as the pixel centers of the reference raster. We organized the resulted covariate matrix in a tiling system to speed up overlay and prediction.

Due to the large number of covariates generated, we analysed their variable importance (VI) to gain insight into the behaviour of carbon soil data and interpretability of covariates influence in its modelling. The selection of the most relevant covariates improves the model accuracy, avoiding the risk of variable redundancy and overfitting (Gregorutti *et. al*., 2016). We used three VI approaches with different level of complexity to select the final covariates considering different relationship functions between target variable and predictors. These variable selection techniques, also called feature selection or subset selection, were based on linear, Bayesians analysis and partial dependence plot (PDP) methods. We implemented the VI approaches in R software ('vip', 'rethinking' and pdp' packages, respectively). The use of model-specific VI scores (such as linear models) are not necessarily comparable across different type of models. We also applied model-agnostic methods (Bayesians and PDPs methods) to compare VI values based on learning algorithm. For the multilinear relationship, the absolute value of the t-statistic was used as a measure of VI. In the Bayesian data analysis, we fitted the probability model to data by specification of distributional assumptions such as likelihood, parameters, and priors. The model was fitted using Markov Chain Monte Carlo (MCMC) sampling. The resulting estimates, conditional on the data, are known as the posterior distribution. This conditioning is governed by the rules of probability theory, which defines a uniquely logical posterior for every prior, likelihood, and data (McElreath, 2016). Comparing with ordinary linear regression, the Bayesian data analysis are independent of the sample data size and structure, thus a more robust simulations are generated for more contemporary big data set. Based on the best combination of the distributional assumptions, a model was built with the iterative selection of more significant covariates. The final covariates selection were related to the model with the best information criteria, based on Akaike information criteria (AIC), the effective number of simulated sample (>= 300), fewer number of covariates and lower standard error value. The PDPs model-agnostic approach is based on quantifying the "flatness" of the PDPs of each covariates to graphically display the effect of the covariate space on the estimated prediction surface. We fitted a projection pursuit regression model and constructed PDPs for each feature. Finally, we ranked and scored the predictors in terms of their relative influence on the response.

The final covariates data set were formed by the higher relative influence covariates common to the three feature selection techniques. Since the relationship between soil properties and environmental variables are usually complex, the results of PDP models -based on non-linear relationship- were prioritized over the others, where appropriate. Prior to model building, a regression matrix was performed including this final selection of covariates. The methodological scheme for the regression matrix was similar to the covariate stack in order to reduce geometric distortions.

Due to the complex relationship between SOC and environmental variables, and frequently non- linear, we combined a multi-model ensemble method to predict the estimation of SOC spatial variability and its associated uncertainty (Shangguan *et al*., 2017; Wang *et al*., 2018). We used the following three modelling approaches to predict SOC:

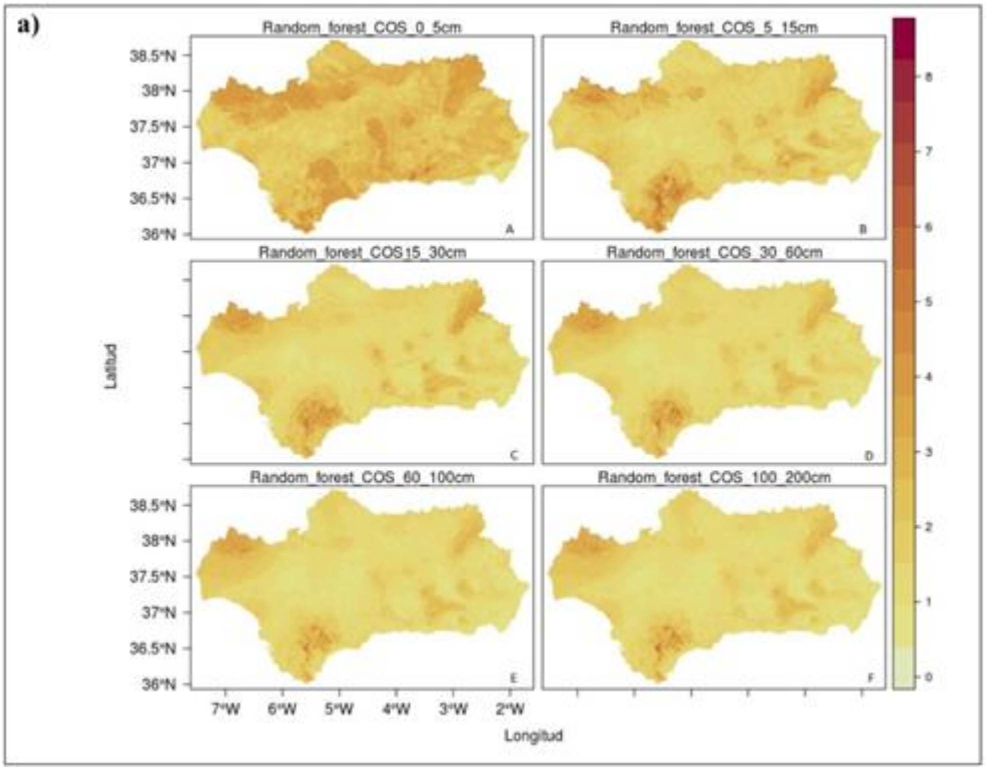
- Quantile regression forest (QRF): QRF is a generalization of random forest (RF) used to estimate an accurate approximation of the full conditional distribution of the response variable for each pixel value. We used qrf algorithm implemented in R software environment for statistical computing in packages: quantregForest (QRF) (Meinshausen, 2006). The QRF validation statistics was calculated from out-of-bag error.

- Machine learning models (MLR): a modelling ensemble approach using non-parametric models such as random forest (RF), deep learning (DL), cubist (Cb) and weighted k-nearest neighbor classifier (kknn). We created an optimal weighted average of those models, also known as an "ensemble", using the test data performance. The machine learning models was fitted using the caret package within R (Kuhn, 2016), using a five fold cross validation with the xy spatial component.

- Auto-machine learning (AutoML): a modelling ensemble of the best models where various steps of the machine learning process are automated such as model selection and hyperparameter optimization of the learning algorithm and featurization. We used H2O package for R which unified a variety of machine learning algorithms. The AutoML object includes a “leaderboard” of models that were trained in the process, including the 10-fold cross-validated model performance. The models are ranked by a default metric.

The information criteria to assess the fit of the different models were determination coefficient (R^2), the root-mean-square error (RMSE) and concordance correlation coefficient (CCC). Each model was evaluated by comparing the predicted SOC values performed with the 75% of the data set, with a validation data set (25% remaining). We repeated this process three time for each model approach to obtain the mean of the models evaluation statistics (CCC, normalized RMSE). We calculated the standard deviation of resulting three maps in each pixel of the spatial prediction. This final map were interpreted as a surrogated of the spatial uncertainty associated with the SOC estimates for each pixel value, and for each modelling approach.

The three maps of the SOC spatial prediction, and their associated uncertainty maps, generated by the different modelling approaches were used for the final maps. The most precise model, i.e. with the lowest uncertainty value, were assigned to each pixel. This process were repeated in each six depth layers to obtain the six SOC density maps and their associated uncertainty maps (Figure 3.67).



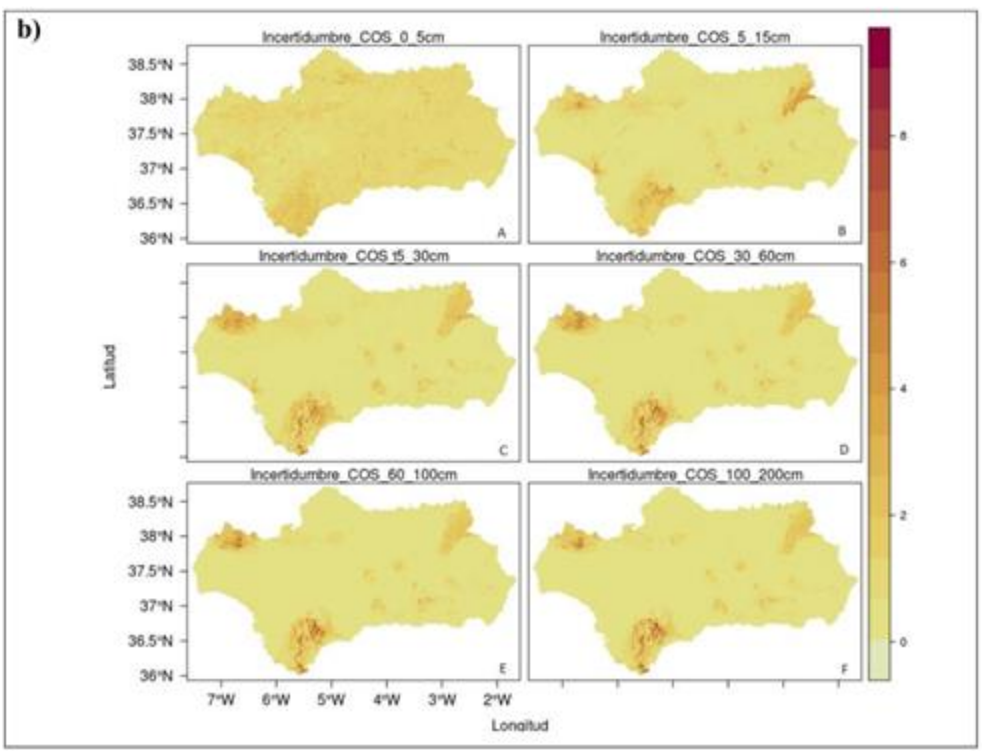


Fig. 3.67: (a) Maps of organic carbon distribution in Andalusia, at different soil depths. Values expressed as %. Explained variance between 57% and 63%. (b) Maps of uncertainty distribution associated with organic carbon

estimation in Andalusia, at different soil depths. Values expressed as %. These maps show the uncertainty for a 95% confidence interval.

Key results

Benchmark map of SOC density across six depth layers: mean values (%) and associated uncertainties spatially explicit (%).

Innovative impact

Satellite-derived indices MODIS indices derived from Google Earth Engine to improve SOC. The SOC map methodology tested in this study can be applied over large areas, such as the whole country, to gain insights into national SOC profile, and subsequently into carbon stock, which remains a key challenge still unsolved in Spain.

Related Copernicus domains: Land Monitoring and climate change.

Area of interest

Although the methodology can be apply for national purposes thanks to the availability of the data, we focused our study case in Andalusian Region. This Mediterranean region is located in south of Spain Peninsula and covers an area of about 87 000 km2. The area is characterized by a typically Mediterranean climate, a complex topography (ranging from 0 to 3479 m), cambisols and regosols as the main soil types, and a larger land cover of natural areas (51% approx.) versus croplands (44%).

Algorithm and data used

Soil organic carbon content (%); PP: Precipitation (mm), Tmax: Annual Mean Maximum Temperature (oC), Tmin: Annual Mean Minimum Temperature (oC), Tmed: Annual Mean Temperature (oC), DEM: Digital Elevation Model, AH: Analytical Hillshading, S: Slope AS: Aspect, PlC: Plan Curvature, PC: Profile Curvature, LC: Longitudinal Curvature, CI: Convergence index, CD: Closed Depressions, TCA: Total Catchment Area, TWI: Topographic wetness index, LsF: Ls Factor, CNBL: Channel network base level, VDCN: Vertical Distance to Channel Network, VD: Valley depth, RSP: Relative slope position, MRVBF: Multiresolution valley bottom flatness index, MRRTF: Multiresolution ridge top flatness index.

Satellite-derived indices: the mean, maximum, minimum, standard deviation, coefficient of variation, range, relative range, sine and cosine of the dates of maximum and minimum of the NDVI: Normalized Difference Vegetation Index; EVI: Enhanced Vegetation Index.; NDWI: Normalized Difference Water Index; Albedo; LST: Land Surface Temperature.

Relevant literature

Alcaraz-Segura D, Lomba A. Sousa-Silva R, Nieto-Lugilde D, Alves P, Georges D, Vicente JR, et al. Potential of satellite-derived ecosystem functional attributes to anticipate species range shifts.

International Journal of Applied Earth Observation and Geoinformation 2017; 57:86±92. https://doi.org/10.1016/j.jag.2016.12.009

Amatulli, G., McInerney, D., Sethi, T. et al. Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. Sci Data 7, 162 (2020). https://doi.org/10.1038/s41597-020-0479-6

Fathololoumi, S., Vaezi, A.R., Alavipanah, S.K., Ghorbani, A., Saurette, D., Biswas, A., 2020. Improved digital soil mapping with multitemporal remotely sensed satellite data fusion: A case study in Iran. Science of The Total Environment 721, 137703. https://doi.org/10.1016/j.scitotenv.2020.137703

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetaryscale geospatial analysis for everyone. Remote Sensing of Environment 2017. https://doi.org/10.1016/j.rse.2017.06.031

Gregorutti, B., Michel, B. & Saint-Pierre, P. Correlation and variable importance in random forests. Stat Comput 27, 659–678 (2017). https://doi.org/10.1007/s11222-016-9646-1

Huete AR, Liu HQ, Batchily K, van Leeuwen W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sensing of Environment 1997; 59(3): 440±451. https://doi.org/10.1016/S0034-4257(96)00112-5

Jax K. Ecosystem Functioning. Cambridge: Cambridge University Press; 2010.37. Cabello J, FernaÂndez N, Alcaraz-Segura D, Oyonarte C, Piñeiro G, Altesor A, et al. The ecosystem functioning dimension in conservation: insights from remote sensing. Biodiversity and Conservation 2012; 21(13):3287±3305. https://doi.org/10.1007/s10531-012-0370-7

Kuhn, M. The caret package. R Foundation for Statistical Computing, Vienna, Austria. URL https://cran. r-project. org/package= caret 2012.

McElreath, R. (2016). Statistical Rethinking: A Bayesian Course with Examples in R and Stan (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315372495

Meinshausen, N., 2006. Quantile Regression Forests. Journal of Machine Learning Research 7, 983– 999.

Paruelo JM, JobbaÂgy EG, Sala OE. Current Distribution of Ecosystem Functional Types in Temperate South America. Ecosystems 2001; 4(7):683±698. https://doi.org/10.1007/s10021-001-0037-9

Peng, L., Cheng Zhi, Q., Xing, Z.A., Zhi Wei, H., Nai Qing, F., Yi Jie, W., 2020. A case-based method of selecting covariates for digital soil mapping. Journal of Integrative Agriculture 19, 2127–2136. https://doi.org/10.1016/S2095-3119(19)62857-1

Potter C, Gross P, Genovese V, Smith M- L. Net primary productivity of forest stands in New Hampshire estimated from Landsat and MODIS satellite data. Carbon Balance and Management 2007; 2 (9). https://doi.org/10.1186/1750-0680-2-9 PMID: 17941989

Schillaci, C., Lombardo, L., Saia, S., Fantappiè, M., Märker, M., Acutis, M., 2017. Modelling the topsoil carbon stock of agricultural lands with the Stochastic Gradient Treeboost in a semi-arid Mediterranean region. Geoderma 286, 35–45.https://doi.org/10.1016/j.geoderma.2016.10.019

Yigini, Y., Olmedo, G.F., Reiter, S., Baritz, R., Viatkin, K., Vargas, R., 2018. Soil organic carbon mapping: Cookbook 2nd Edition. FAO.

Success story #6 for LULUCF report in Spain. Mapping grasslands at a regional scale: from remote sensing to land use management and productivity.

Application field: Mapping LULUCF grassland classes.

Abstract

In this work, we applied a framework of predictive mapping based on environmental variables and time series of remote sensing data to estimate grassland distributions, one of the most poorly mapped natural systems in the whole of Spain. The method is intended to provide fine scale maps large areas. In a first step, the classification is made spatially explicit on the basis of an available grassland land use layer derived a 1 meter of spatial resolution for the whole region of Cantabria. Then, by means of in situ data of different EUNIS categories o in Cantabria, we proceeded mapping grasslands typologies according to their land use regime: natural and managed (by livestock and mowing) based on the values of intra-annual oscillation of the NDVI index obtained from S2 imagery since 2017 to 2021. In addition, the relationship between these categories and the productivity of grasslands was studied according to the PUERTO model.

Introduction

Grasslands (GL) are one of the most poorly mapped natural systems in the whole of Spain. Their management and conservation status in rural areas, where extensive livestock farming represent one of the main economic activities, needs remain in an ecologically sustainable and productive state over time. Their progressive abandonment due to the rural exodus means that classic formations such as communal pastures are being lost or degraded, i.e. invaded first by scrub and then even by woody species in a characteristic process of natural succession.

At a geographical scale, traditional techniques based on the photointerpretation of aerial images with the support of field campaigns are neither efficient nor cost-effective to objectively monitor management practices in these widely distributed and dynamic complex systems. The development of cutting-edge technology programmes such as Copernicus in Europe provides public access to in situ and satellite datasets that enable the development of geographic-scale mapping products with high spatial and thematic detail. The exploitation of these data provides valuable evidence of the success of this European initiative. The results obtained will serve as an example for future satellite missions (e.g. FLEX, PRISMA, HyspIRI, etc.) related to vegetation fluorescence or hyperspectral resolutions, useful in variable and complex land cover areas.

In order to meet the needs of level 2 data for the estimation of E/A in GL and to advance in the knowledge of their distribution, change processes and ecological dynamics, a mapping and monitoring methodology is defined. This methodology is based on the use of remote sensing techniques and the *in situ* knowledge on the different GL typologies that can be used for classifying these complex systems by land use and management. In an exploratory exercise, with the aim of advancing knowledge on their distribution and ecological dynamics, the grassland mapping methodology has been applied in Cantabria using ground truth points defined by botanical fieldwork and remote sensing data that are complemented with the results of the PUERTO model.

To achieve this goal, this case study pursues a number of specific objectives: 1) Identifying typologies of grasslands based on management, use and biomass estimations that enhance the knowledge and accuracy of the estimations: natural and managed (by livestock and mowing); 2) Mapping these grasslands typologies by applying methods based on remote sensing technologies; 3) Developing tools for detection and monitoring of annual changes in land use and/or use intensity, management and biomass gains and losses and applying those tools temporarily. With this mapping objectives achieved, we will be also able to: 4) Evaluating carbon emissions and removals from due to changes within the grasslands and set the procedure to update that estimations for the 2021-2030 period; 5) Estimating differences in carbon emissions and removals derived from transitions among grassland categories for upcoming years; 6) Generating a carbon emissions and removals estimation system for Spain and extendible to other European countries.

Main expected result will end up with an improved method of the specific LULUCF category grassland mapping and further carbon accounting and reporting at the national level in Spain, extendible to other European countries, as well as estimations and associated uncertainty measures of the carbon emissions and removals for some specific dates. The method will also provide tools for detection and characterization of annual land use and use intensity changes among different grassland classes.

<u>Methodology used/Workflow</u>

A first step corresponded to a spatially explicit approximation of the area occupied by grassland in the region of Cantabria (Figure 3.68). It has been obtained by means of a supervised classifications made by high resolution vegetation models obtained from LIDAR data at 1 meter of spatial resolution, time series PNOA NDVI and the Spanish Forestry Maps at a scale 1:25000. This allowed defining all land cover polygons with a very high suitability of being GL communities.

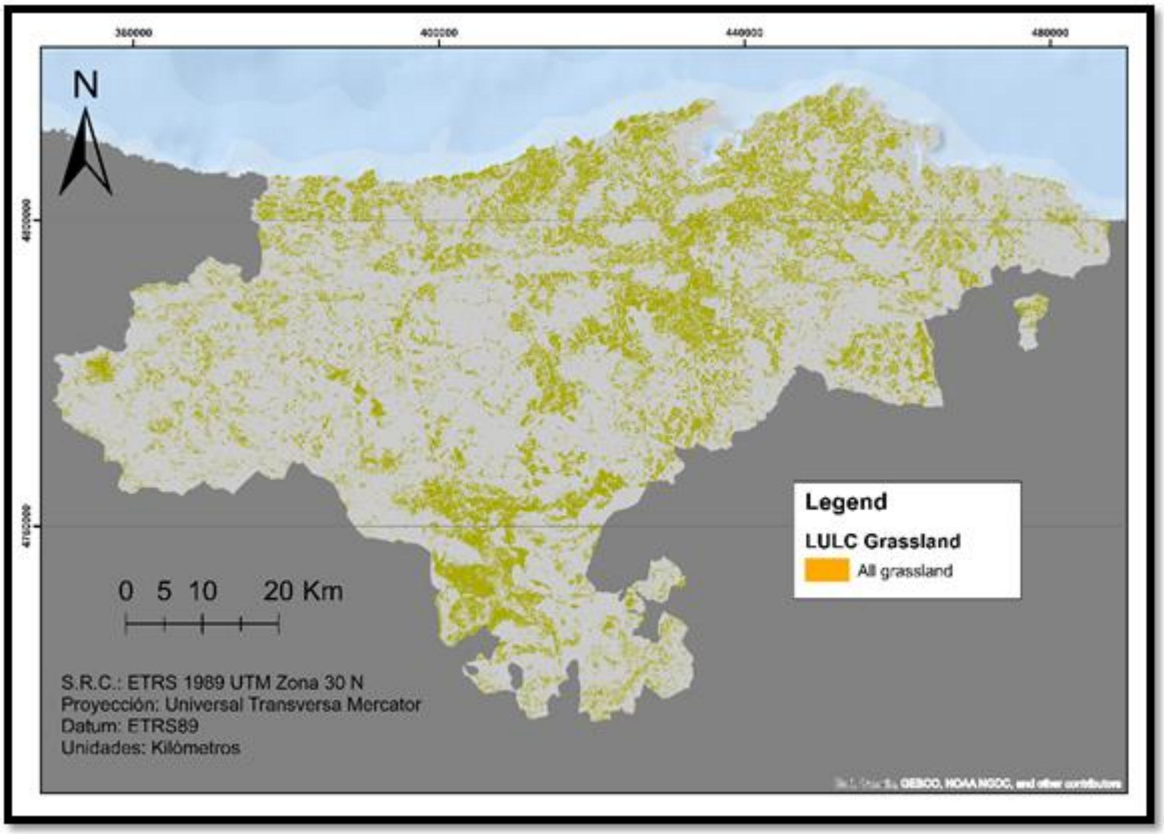


Fig. 3.68. Grassland area for the whole territory of Cantabria.

On the other hand, based on 3,596 points of different EUNIS categories (Table 3.15), defined as grassland by botanists fieldwork cross the entire region, we obtained the following categories:

- Natural grassland, comprising the following EUNIS categories: E1.11; E1.26; E1.31; E1.53; E1.71; E1.72; E3.41; E3.42; E3.1X; E3.52; E4.31; E4.3X; E4.3Y; E4.41; E4.42; E4.43; E5.2X; E5.2Y; E5.43 and E5.53

- Managed or disturbed grassland, differentiating in these between managed by livestock action (categories E2.11 and E2.21) and managed by mowing (category E2.23).

Table 3.15. List of EUNIS 4 typologies for grassland and peatland with name (English and Spanish) and correspondence (if any) with Annex I of the Habitat Directive (HD).

EUNIS4	**Name(EN)**	**Nombre(ES)**	**HD**
E111	Euro-Siberian rock debris swards	Prados calcáreos de suculentas	6110

E126	Sub-Atlantic semi-dry calcareous grassland	Lastonares calcícolas	6210
E131	West Mediterranean grassland	Prados xerofíticos submediterráneos	6220
E153	Iberian fescue frost-influenced grassland	Prados crioturbados montanos	No
E171	Nardus stricta swards	Cervunales acidófilos atlánticos	6230*
E172	Agrostis-Festuca grasslands	Prados mesófilos atlánticos	No
E211	Unbroken pastures	Prados de manejo intenso	No
E221	Atlantic hay meadows	Praderas de heno del Atlántico	6510
E223	Medio-European submontane hay meadows	Prados de siega	6510
E341	Atlantic and subatlantic humid meadows	Pastos húmedos atlánticos	No
E342	Juncus acutiflorus meadows	Prados-juncales cantábricos	No
E31X	Submediterranean Holoschoenion	Prados húmedos mediterráneos	6420
E351	[Molinia caerulea] meadows and related communities	Prados-juncales acidófilos	6410
E352	Heath Juncus meadows and humid Nardus stricta swards	Cervunales higro-turbosos	No
E431	Alpic [Nardus stricta] swards and related communities	Cervunales orocantábricos	6230

E43X	Orocantabrian acidophilous stripped grasslands	Prados alpinos silicícolas	6160
E43Y	Orocantabrian acidophilous Festuca eskia grasslands	Prados alpinos con Festuca eskia	6140
E441	Closed calciphile alpine grassland	Prados alpinos calcícolas continuos	6170
E442	Wind edge [Kobresia myosuroides] swards	Prados alpinos con Kobresia myosuroides	6170
E443	Calciphilous stepped and garland grassland	Prados alpinos calcícolas crioturbados	6170
E52X	Thermophilous forest fringe of base-rich soils	Bordes forestales basófilos	No
E52Y	Cantabrian forest fringe of acidic nutrient-poor soils	Bordes forestales acidófilos	No
E543	Shady woodland edge fringes	Megaforbios de bosques umbrosos	6430
E553	Pyreneo-Iberian tall-herb communities	Megaforbios subalpinos	6430

Not all points correspond to the same quality of data. Some have been collected during field campaigns with in situ work (high quality) while others are derived from photo-interpretation and cross-checking with different inventories (lower quality). The distribution of these points is shown below (Figure 3.69):

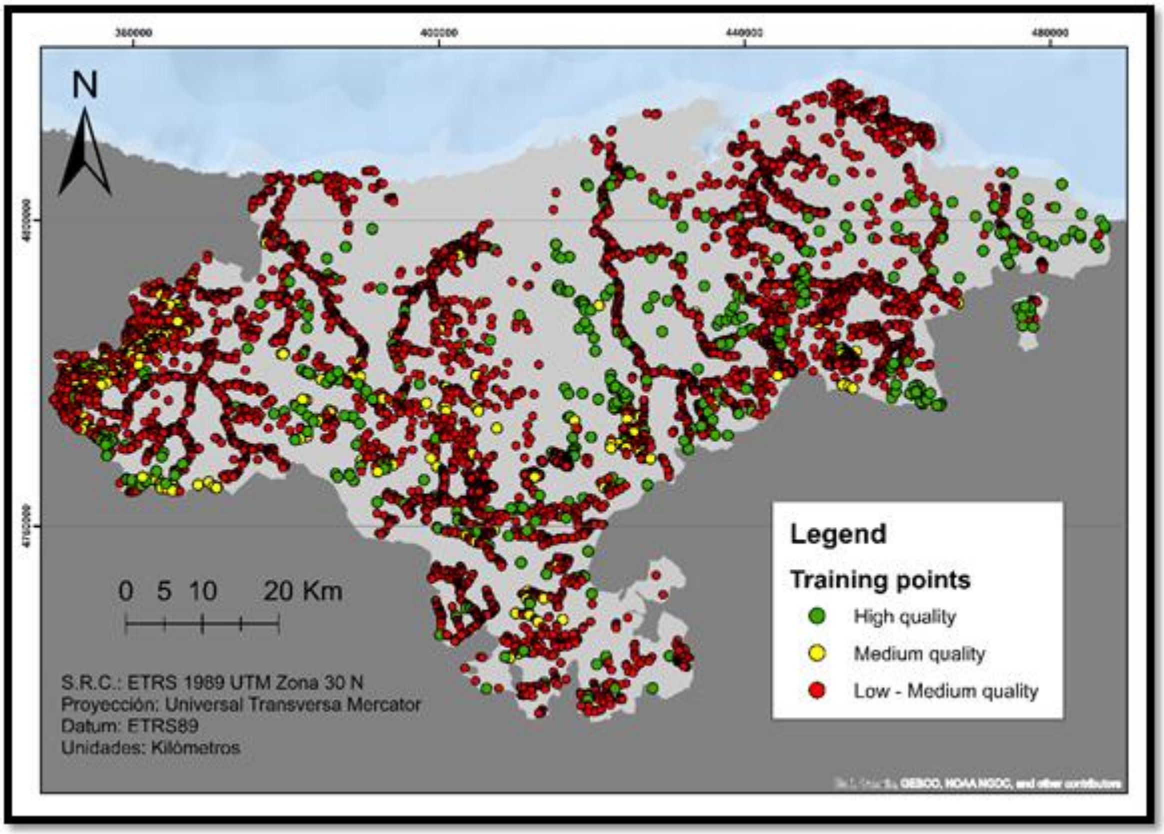


Fig. 3.69. Distribution of points used in the process and their quality.

For GL mapping purposes, the region was split bioclimatic floors, being the “Colino” between 0 and 500 metres above sea level, “Montano” between 500 and 1,500 metres above sea level and “subalpino-alpino” up to 2,650 metres above sea level. In the latter, there are no categorised mowing points, so they all correspond to livestock farming.

In turn, we developed a semi-automated procedure for downloading, correcting and processing all S2 available imagery of the study area. We calculated, among others, the NDVI indices for all S2 images and obtained a raster of NDVI inter- and intra-annual variation with values between 0 and 1, where lower values indicate lower variability and higher values indicate higher variation of this index across time, which is associated with greater phenological (or biomass) changes related to natural and/or anthropogenic disturbances and dynamics (Figure 3.70).

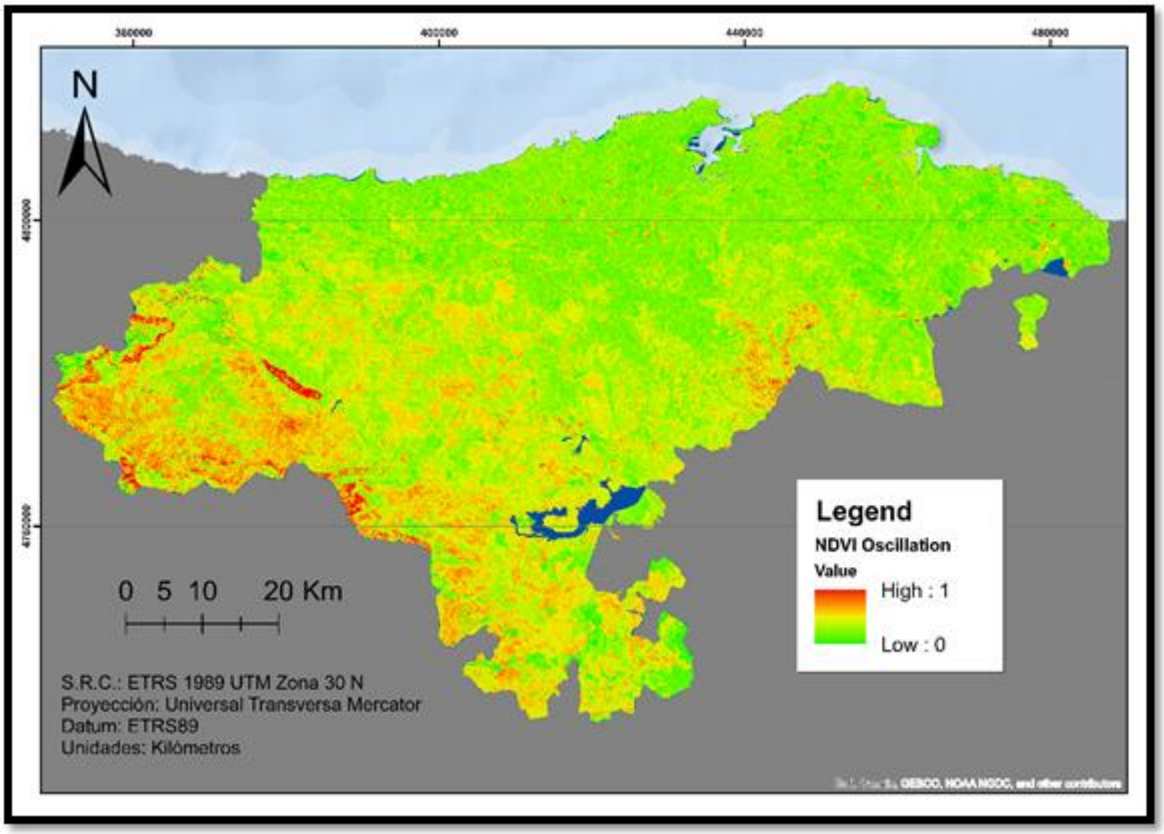


Fig. 3.70. NDVI oscillation values between 0 and 1 for the entire territory.

Next, we obtained the value of the intra-annual variation of the NDVI for each of the EUNIS points and applied ANOVA-tests to look for significant differences among GL groups and topographic strata, obtaining obtaining the following results (Table 3.16):

Table 3.16. Mean NDVI oscillation values for each grass typology. Those whose difference is statistically significant according to the ANOVA test are marked with (*).

Bioclimatic floor	Grassland typology		Mean NDVI oscillation value	
Colino	Natural		0.16834*	
	disturbed grass	mowing	0.23170*	0.22630*

		livestock		0.23190*
Montano	Natural		0.29704*	
	disturbed grass	mowing	0.34620*	0.37870*
		livestock		0.31920*
Subalpino-alpino	Natural		0.40090	
	disturbed grass (livestock)		0.48300	

Based on these values, we classified the grassland LULC vector layer (i.e. all polygons corresponding to GL category) by assigning to each polygon (Figure 3.68) the mean NDVI variation value of all pixels within it. In addition, the average NDVI values at the end of August were extracted for each polygon (Figure 3.71), as well as the values from the results of the PUERTO grassland productivity model (Figures 3.72), scaled from 0 to 1. We obtained the results summarized in the Table 3.17.

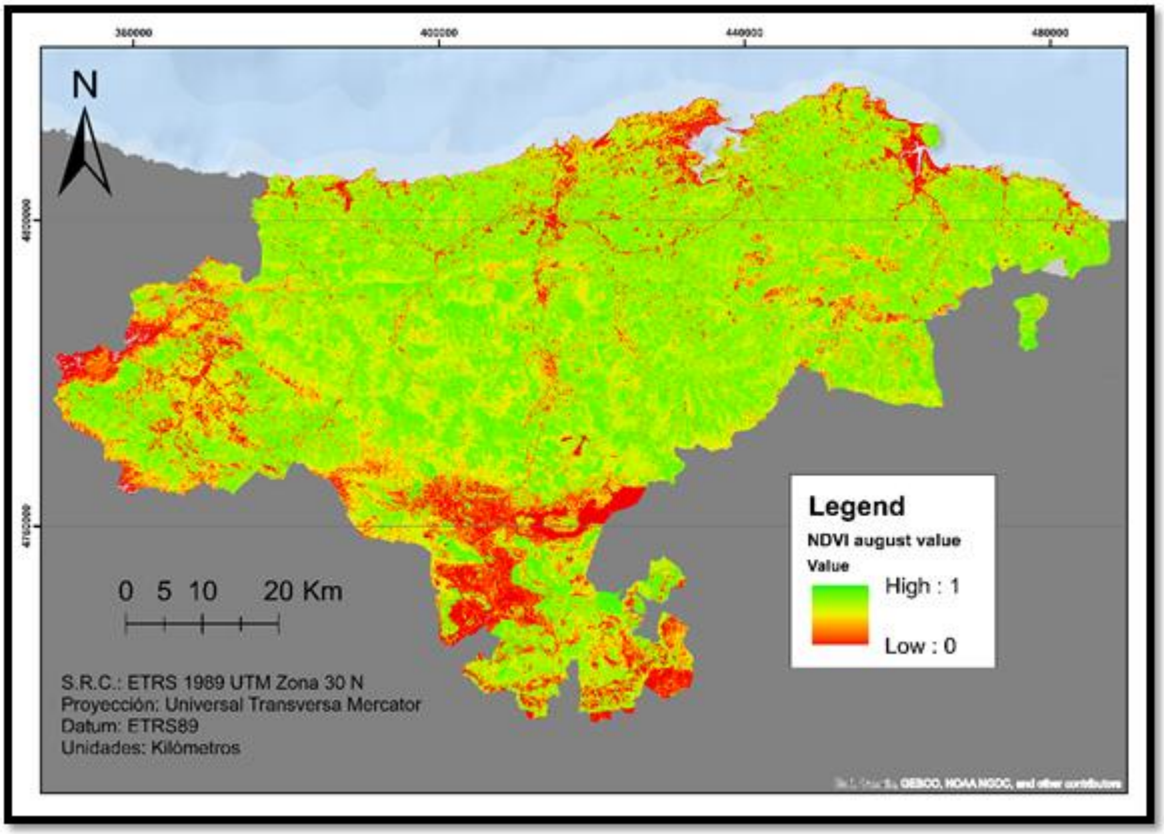


Fig 3.71. NDVI values in August for the whole territory.

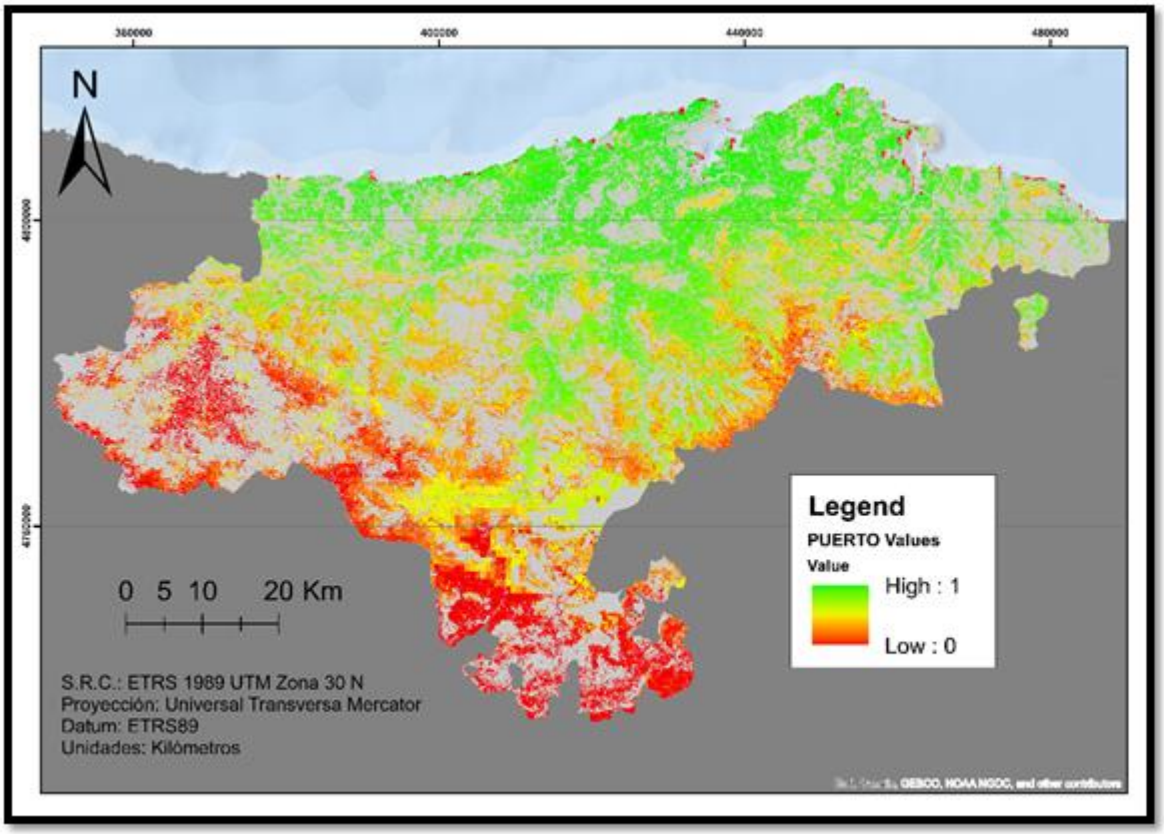


Fig 3.72. Values of the territory-wide PUERTO productivity model.

Table 3.17. Results of the the average NDVI values at the end of August extracted for each polygon, as well as the values scaled from 0 to 1 from the results of the "PUERTO" grassland productivity model.

<table>
<tr><th>Bioclimatic floor</th><th colspan="2">Grassland typology</th><th colspan="2">August NDVI value</th><th colspan="2">PUERTO value</th></tr>
<tr><td rowspan="3">Colino</td><td colspan="2">Natural</td><td colspan="2">0,469267</td><td colspan="2">0,724221</td></tr>
<tr><td rowspan="2">disturbed grass</td><td>mowing</td><td rowspan="2">0,470261</td><td>0,469553</td><td rowspan="2">0,838499</td><td>0,850613</td></tr>
<tr><td>livestock</td><td>0,470539</td><td>0,834068</td></tr>
<tr><td>Montano</td><td colspan="2">Natural</td><td colspan="2">0,483394</td><td colspan="2">0,434023</td></tr>
</table>

	disturbed grass	mowing	0,44194	0,445353	0,459883	0,433544
		livestock		0,455190		0,415914
Subalpino-alpino	Natural		0,400784		0,186367	
	disturbed grass (livestock)		0,39658		0,185262	

The PUERTO model simulates the functioning in space and time of the most relevant processes of the climate - soil - grass - herbivore - management system that make up pastoral agro-ecosystems. It is a tool for defining actions on the territory and on livestock, prioritising the types of vegetation to be treated, deciding on working methods, adapting the scope of grazing or promoting the management of other complementary livestock species. In addition, thanks to the existence of highly accurate climatological information at detailed spatio-temporal scales and the advances in the field of geostatistics made possible by current GIS, the model predicts the productivity of grazing areas, taking into account the effect of altitude and orography on this factor.

Results

Through this workflow, 65,522 ha of natural pasture area and 57,975 ha of anthropised pasture area have been obtained, which in turn can be divided into 25,722 ha of mowed area and 32,251 ha of livestock pasture (Figure 3.73 and 3.74). Numbers are in agreement with official statistics of the region, in general terms, and the spatial patterns were accurate in terms of in situ EUNIS data and expert knowledge of the region (Figure 3.75).

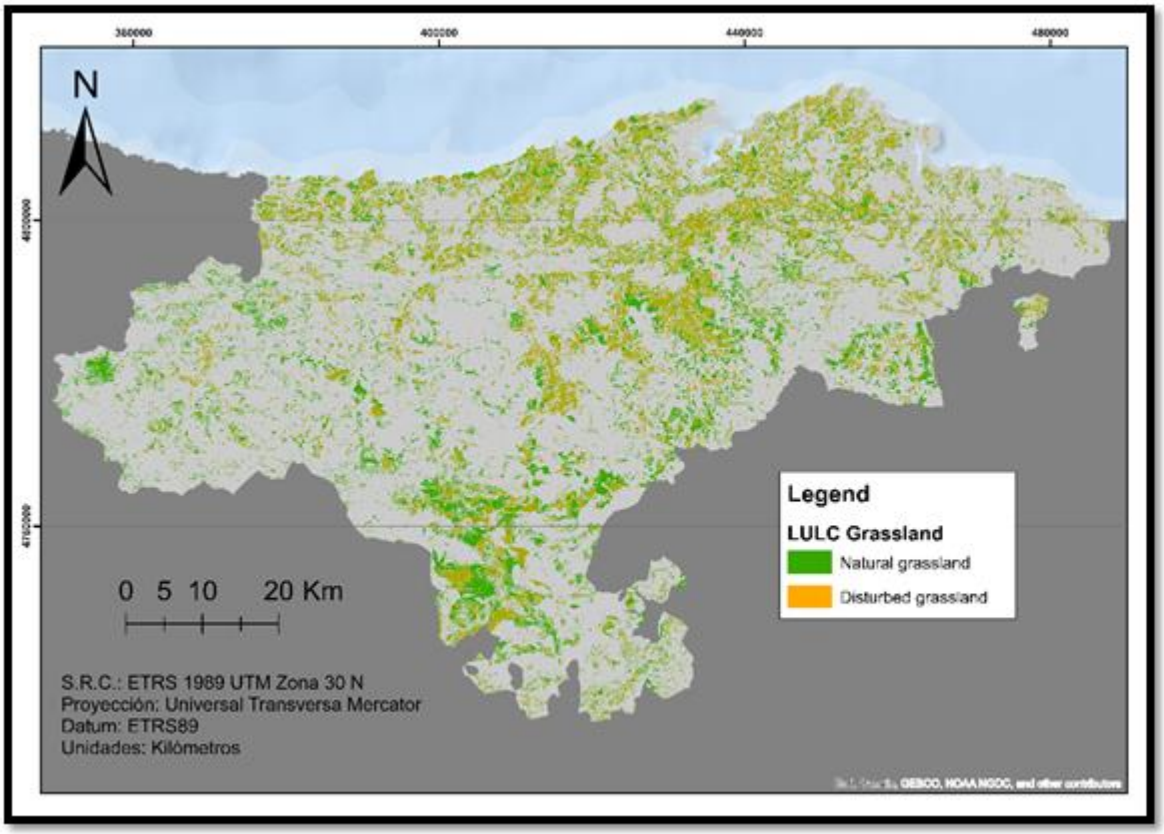


Fig. 3.73. Spatial distribution of natural and disturbed grasslands for the whole territory.

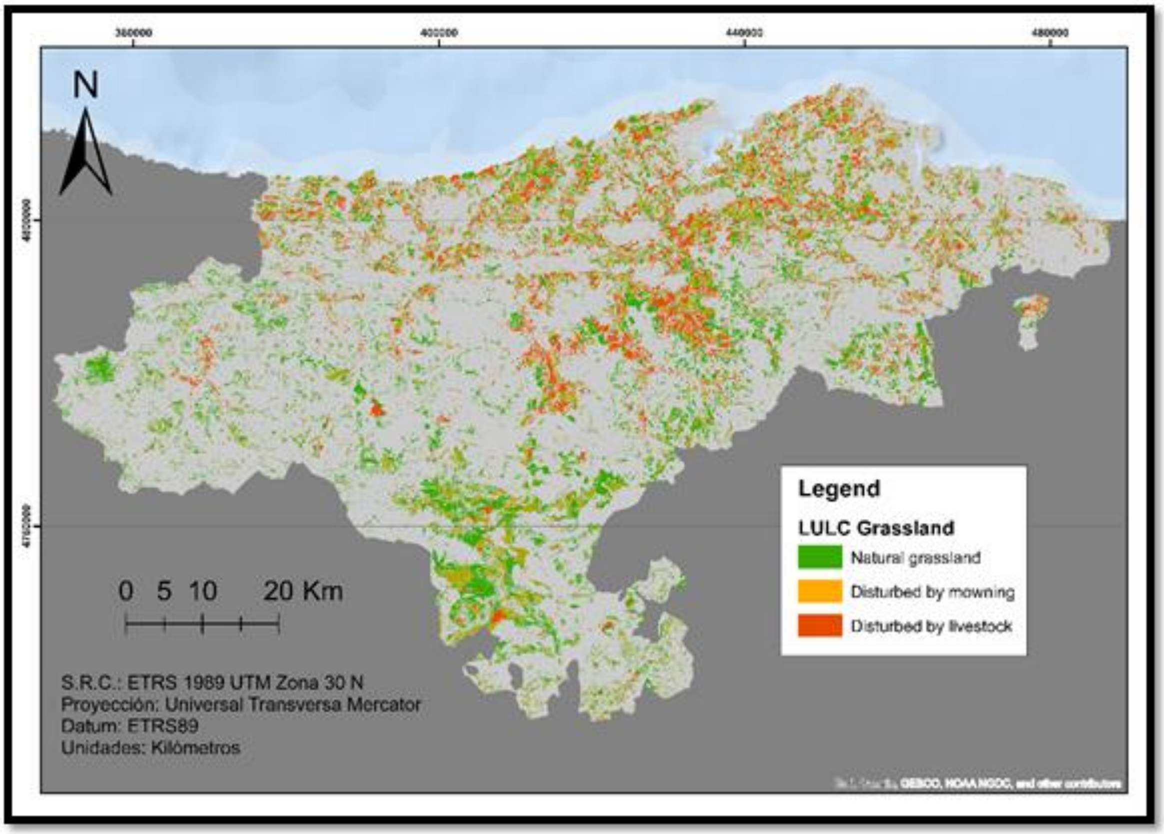


Fig. 3.74. Spatial distribution of natural grassland and grassland disturbed by livestock and mowing for the whole territory.

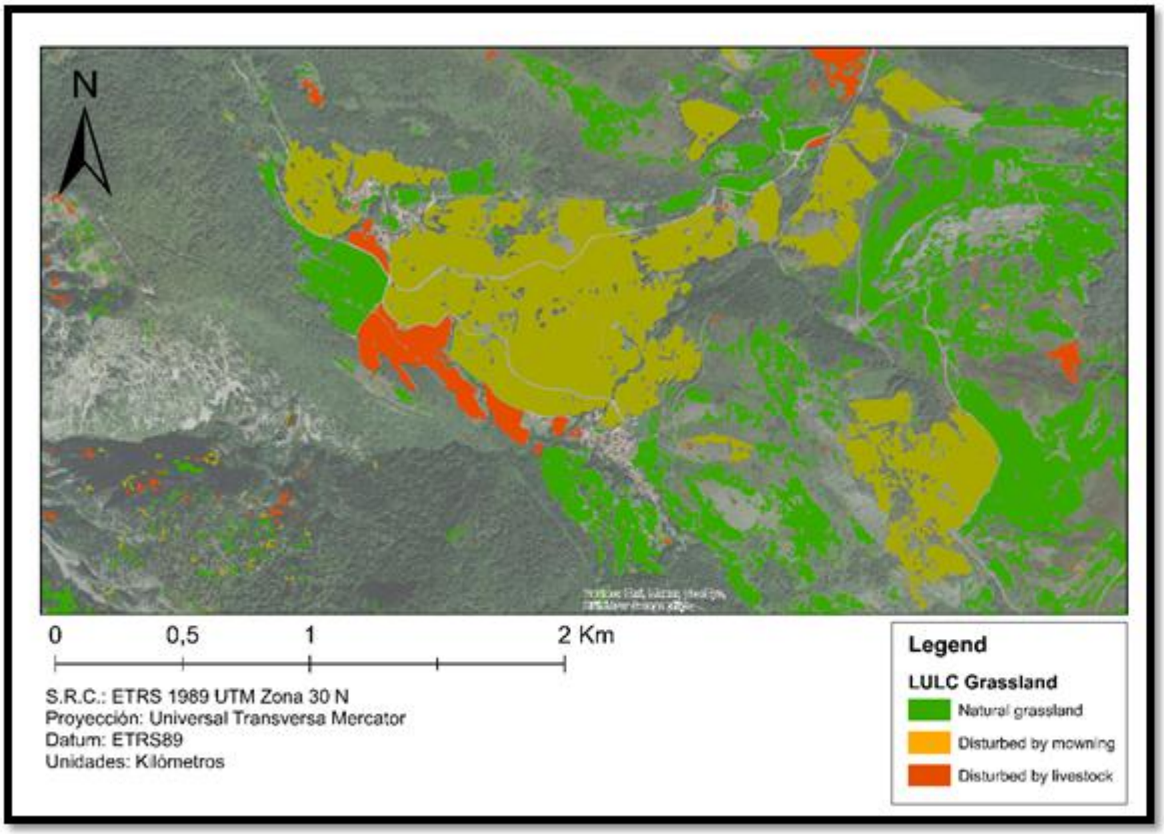


Fig 3.75. Detail of the distribution of pasture typology in Cicera, in Deva river basin.

Natural pastures in the 3 bioclimatic zones showed a lower NDVI variation than those modified by human action because the phenology of the former is based on natural cycles while that of the modified ones is drastically altered in short periods of time by human management of livestock grazing across large areas.

The PUERTO model indicates a higher productivity in the human managed pastures than in the natural pastures for the hill and mountain bioclimatic floors, which corresponds to the historical use of the soils and that the most productive soils are dedicated to mowing.

One step beyond, once determined the capability of the workflow for GL mapping (i.e. to be able differentiating between managed grassland (for grazing and mowing) and natural grassland according to EUNIS typologies, we proceed to obtaining two additional useful indicators for GL mapping (ongoing work):

- the spectrophenological curves from the time series of S2 imagery from which we obtained the NDVI oscillation,

- the methodology for collecting the hyperspectral signatures of all EUNIS GL communities identified in Cantabria.

Table 3.18 and Figure 3.76 show the results of the spectrophenological curve for mowing meadows obtained with time series of Sentinel 2 satellite images (from 2016 to 2020) to study the spectrophenological variability of different grassland types with EUNIS classification collected through in-situ sampling used in former supervised analysis. Table 3.19 and Figure 3.77 show the same results for a EUNIS type of natural grasslands, atlantic acidophilous cervunales (*Nardus stricta*), which results allow us to advance in the definition of specific indicators related to their distribution, structure and functional variables such as productivity or phenological dynamics.

Table 3.18. Phenological interpretation of curve indicators for mowing meadows. Obtained by ITD Medioambiente SL.

<table>
<tr><th colspan="4">Mowing meadows</th></tr>
<tr><th colspan="2">indicator or factor</th><th rowspan="2">Time period</th><th rowspan="2">Value</th></tr>
<tr><th>Acronym</th><th>Phenological interpretation</th></tr>
<tr><td>SOST</td><td>Start of measurable photosynthesis in the plant canopy</td><td rowspan="2">December.</td><td rowspan="2">0.62</td></tr>
<tr><td>SOSN</td><td>Level of photosynthetic activity at the onset of measurable photosynthesis</td></tr>
<tr><td>EOST</td><td>End of measurable canopy photosynthesis</td><td rowspan="2">October</td><td rowspan="2">0.68</td></tr>
<tr><td>EOSN</td><td>Level of photosynthetic activity at the end of measurable photosynthesis</td></tr>
<tr><td>MAXT</td><td>Time of maximum photosynthesis in the plant canopy</td><td rowspan="2">May</td><td rowspan="2">0.80</td></tr>
<tr><td>MAXN</td><td>Maximum level of photosynthetic activity in the canopy</td></tr>
<tr><td>DUR</td><td>Duration of photosynthetic activity (growing season)</td><td>10 months</td><td>ND</td></tr>
<tr><td>AMP</td><td>Peak increase in canopy photosynthetic activity above baseline</td><td>ND</td><td>0.18</td></tr>
<tr><td>TIN</td><td>Canopy photosynthetic activity throughout the growing season</td><td>ND</td><td>7.08</td></tr>
</table>

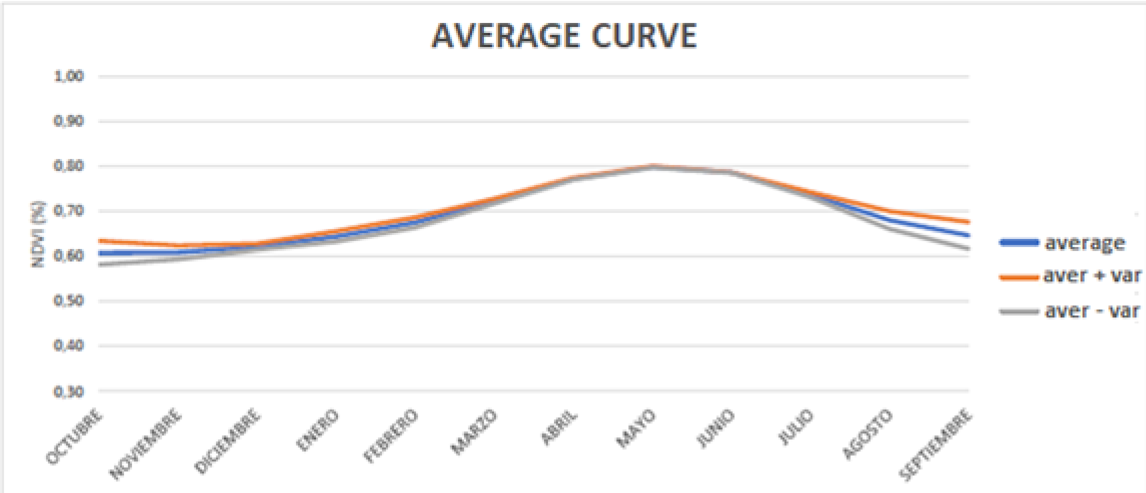


Fig. 3.76. Spectrophenological curve for mowing meadows.

Table 3.19. Phenological interpretation of curve indicators for atlantic acidophilous cervunales. Obtained by ITD Medioambiente SL.

Atlantic acidophilous *cervunales* (*Nardus stricta*)			
indicator or factor		Time period	Value
Acronym	Phenological interpretation		
SOST	Start of measurable photosynthesis in the plant canopy	March	0.56
SOSN	Level of photosynthetic activity at the onset of measurable photosynthesis		
EOST	End of measurable canopy photosynthesis	January	0.53
EOSN	Level of photosynthetic activity at the end of measurable photosynthesis		

MAXT	Time of maximum photosynthesis in the plant canopy	July	0.75
MAXN	Maximum level of photosynthetic activity in the canopy		
DUR	Duration of photosynthetic activity (growing season)	10 months	ND
AMP	Peak increase in canopy photosynthetic activity above baseline	ND	0.18
TIN	Canopy photosynthetic activity throughout the growing season	ND	6.73

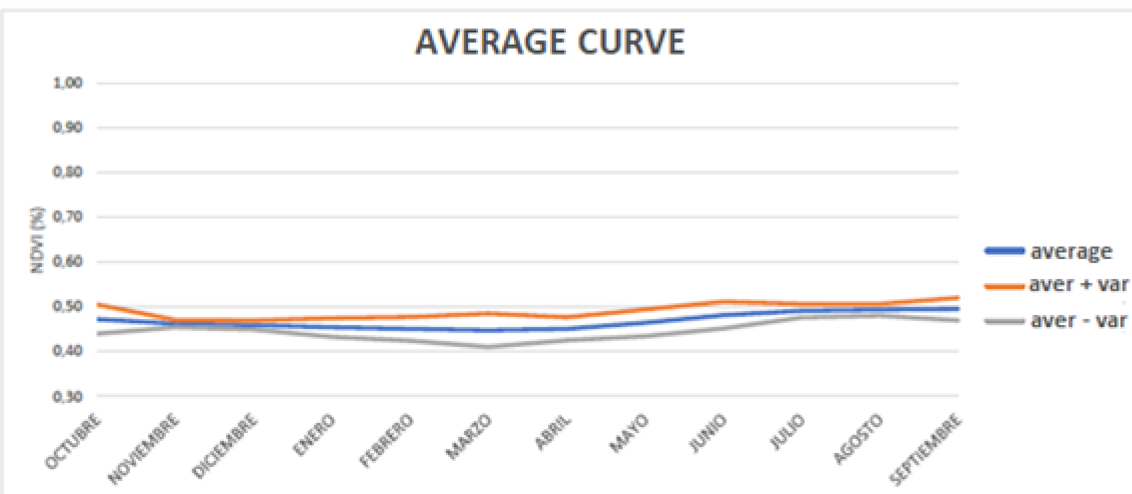


Fig. 3.77. Spectrophenological curve for atlantic acidophilous cervunales.

Based on these curves and as an example, the mean oscillation of the inter-annual NDVI value for different areas of high mountain grasslands (Figure 3.78) and in low valley areas strongly altered by human management (Figure 3.79) is presented below. The results show that high mountain pastures, under limited grazing pressure, show a much smaller inter-annual NDVI oscillation, while those altered by mowing in the lower part of the valley show a large variation, as explained previously.

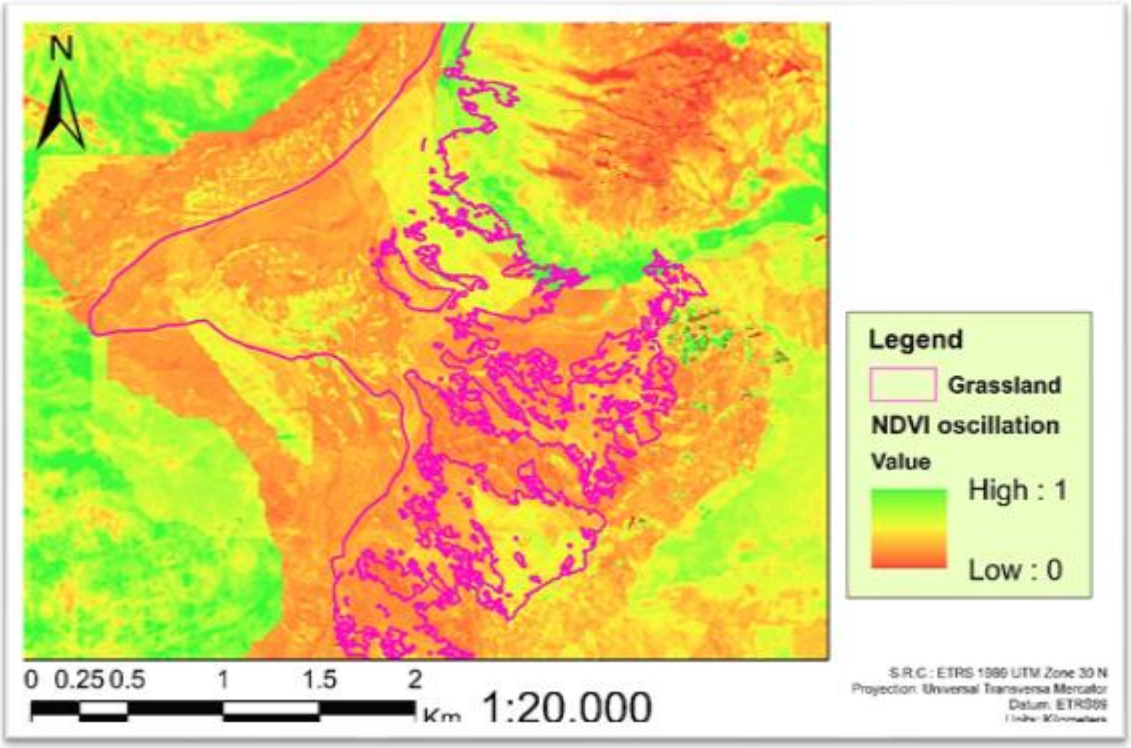


Fig 3.78. NDVI oscillation in high mountain grasslands.

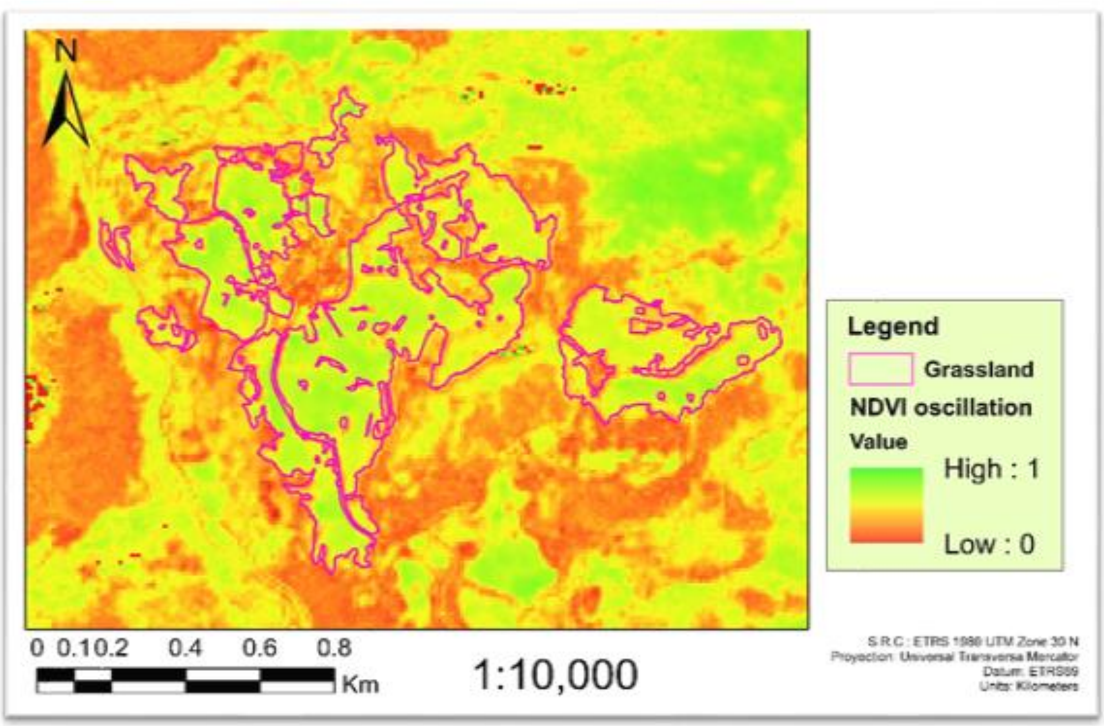


Fig 3.79. NDVI oscillation in valley grasslands, mostly affected by human management.

Once the information available from the spectrophenological curves has been processed (ongoing work), hyperspectral signatures will be obtained for the different types of grassland by means of seasonal sampling (spring, summer, autumn and winter), which will enable the spectrum of habitats in different phenological and physiological states to be characterised. For this purpose, it is proposed to characterise separately a number of defined study areas. Those plots larger than 1 hectare within will be identified, as well as the 30x30m subplots, coinciding with 3 pixels of the MSI sensor of the Sentinel-2A & -2B satellites included within the area of the selected plots. The selection of these pixels will be carried out on the basis of the distribution mask in the geography of the northern peninsular strip of habitats of community interest corresponding to natural and semi-natural grassy formations, and the analysis of the historical time series of the data of those spectral indices of special relevance for this type of formations, derived from satellite sensors. Pixels with a normal reference pattern will be identified, as well as those with an anomalous pattern (systematic deviation). Both groups of pixels will allow both the validation and the calibration (and improvement of the interpretation), respectively, of the data from the satellite sensors, being the object of field sampling. The analysis of the information obtained in situ referring to the spectral resolution of the remote sensors will make it possible to determine which habitats can be differentiated by remote sensing, as well as the capabilities of each satellite for their identification.

In order to replicate the accuracy of the sample measurements at all locations, a standard protocol for the spectral library data collection has been created. The most important thing to consider is that the spectral signature measurements in the field have to be carried out under clear sky conditions (<2/8 clouds and not obscuring the solar disc), preferably during the central hours of the day, between 10:00-15:30 h. The steps to be followed are as showed in Figure 3.80. The steps to follow are:

- Stabilisation of the sensor (once the sensor is switched on, wait 20-30 minutes for the signal to stabilise).

- The collection of metadata associated with the sampling location, atmospheric conditions, lighting conditions, time of day, etc. For each measurement performed, additional information (metadata) on weather conditions, instrumentation and type of sample measured and substrate shall be recorded. In addition, photos of the measurement, the sky, and the sampled study area shall be taken using a standardised procedure. A field identification panel will be used at each sampling site to record sample identifications.

- Once the spectroradiometer is stabilised, the sensor's black current is corrected (this process is usually integrated in the acquisition software).

- The sensor is then calibrated with the reference white panel. For this purpose, the radiance of a white panel (e.g. Spectralon® white 99% from Labsphere) is measured at a certain height (approx. 10-15cm). It is important to note that every time the atmospheric or light conditions change (e.g., clouds, shadows, etc.), the spectroradiometer must be recalibrated (readings of the reference white panel) before obtaining the spectral signatures.

- The next step is to obtain the spectral signatures (radiances and reflectances), with an arrangement of approximately 0o at nadir, and 135o with respect to the solar azimuthal angle. The measurement will be carried out at a distance of approximately 1m above the canopy.

- To ensure consistent and accurate averaging of spectra (minimising the effects caused by illumination differences, viewing angle variations and target material variability), each measured spectrum shall be calculated as the mean value of 10-20 spectra over the same object to ensure an optimal signal-to-noise ratio. In addition, a minimum of five pseudo-replicates must be obtained for each sample, from which the mean spectrum and the uncertainty associated with the reproducibility of the data shall be estimated. In this way, the effect of slight differences in illumination, viewing angles and object variability on the spectrum is minimised. On the other hand, the dark current and reference panel readings for normalisation shall be obtained as the average of 10 and 25 measurements, respectively.

Each target and target panel measurement is derived from an average of multiple samples, and the spectral reflectance (target radiance/panel radiance) is derived by the spectrometer software and stored together with the sample location. All field-based reflectance measurements must be adjusted by a scale factor representing the ratio of a pristine Spectralon panel to field panels.

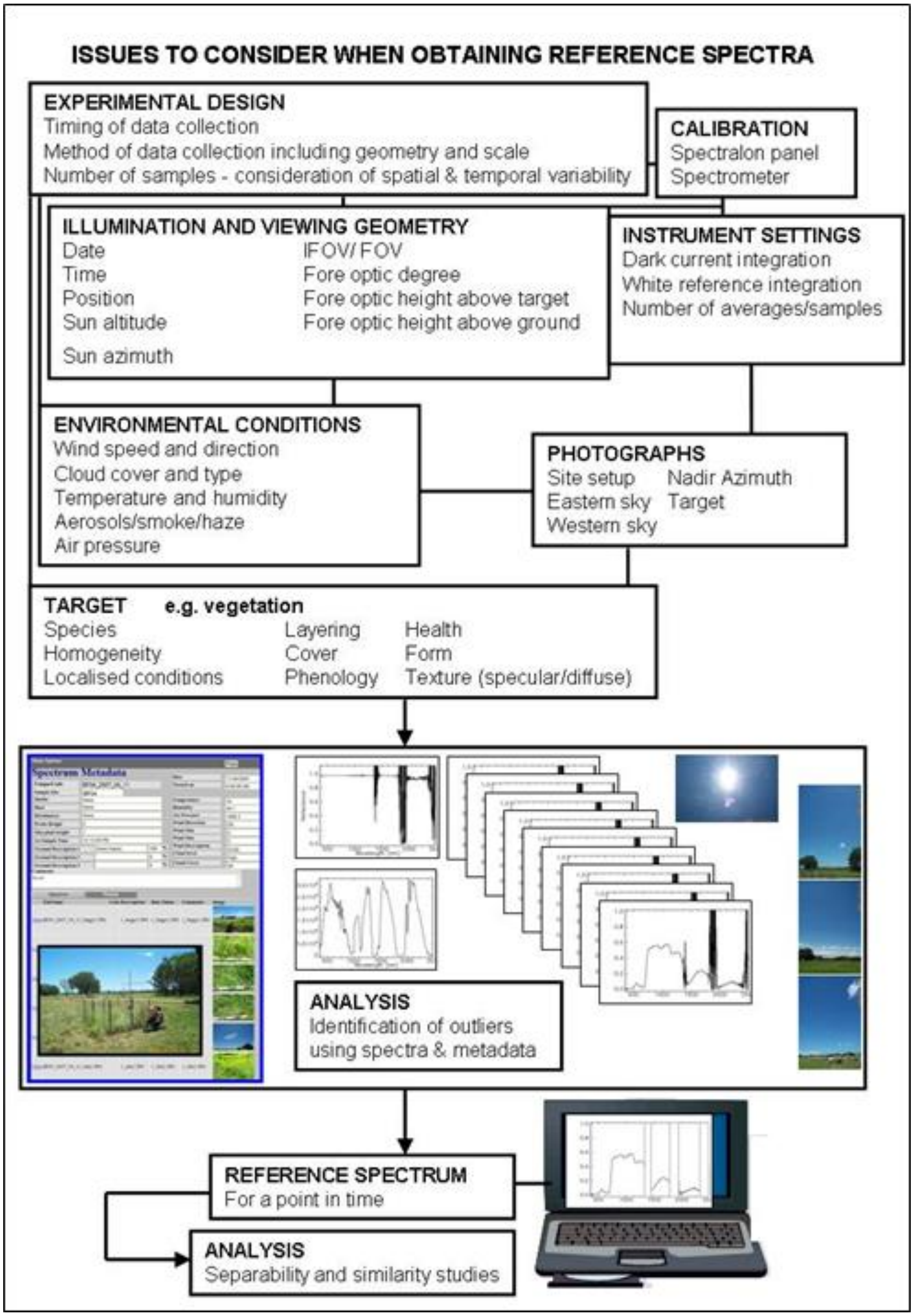


Fig. 3.80. Conceptual diagram of factors affecting spectral.

The processing of the raw spectral reflectance data provided by the spectroradiometer shall consist of the following steps:

- Transformation of relative reflectance to absolute reflectance. The relative reflectance measurements are multiplied by the calibrated reflectance of the reference panel (Spectralon intercalibrated with the pristine laboratory Spectralon) used in the measurement.

- Estimation of the average spectral signatures and standard deviation of the spectra obtained at each measurement point within each sampling station. It is performed from the five pseudo-replicates obtained from each sample under the same conditions. The standard deviation is an estimate of the measurement reproducibility (i.e., uncertainty).

- Disregard of spectral regions that are strongly affected by atmospheric absorption: 1350-1440 nm, 1790-1980 nm, 2360-2500 nm (Hueni, 2006). These regions are ignored in the following steps of the spectral signature processing.

- Elimination of reflectance values corresponding to the ultraviolet region (below 400 nm), given the increased sensor noise in this region.

- Correction of the jump between the reflectance values at the boundary of the domains of the three sensors (one covers the VNIR, Visible-near infrared, and the other two the SWIR, Shortwave infrared) that are integrated in the ASD FieldSpec 4 spectroradiometer. In this case, only the VNIR and SWIR1 transition is corrected, since the SWIR1 and SWIR2 transition occurs in the region of the spectrum affected by atmospheric noise, especially in the data collected in the field. This correction is based on a multiplicative approach, which consists of calculating a correction factor, expressed as the quotient between the reflectance of the first band of the reference sensor (SWIR1) and the reflectance of the last band of the sensor to be corrected (VNIR), and multiplying it by the reflectance of each of the bands of the sensor to be corrected (VNIR). This means that parts of the spectrum with high reflectance values are more affected by the correction than wavelengths with low reflectances.

Finally, to obtain the representative signature of each habitat of community interest corresponding to natural and semi-natural grassland formations, the calculation of the average of the signatures obtained in the different sampling points for that habitat or species in a similar physiological and phenological state will be carried out. The uncertainty shall be estimated as the square root of the sum of the square of the standard deviations of each spectrum.

Relevant literature

Alonso-Martínez, L., Ibañez-Álvarez, M., Brolly, M., Burnside, N. G., Calleja, J. A., Peláez, M., ... & Serrano, E. (2020). Remote mapping of foodscapes using sUAS and a low cost BG-NIR sensor. Science of the Total Environment, 718, 137357.

Álvarez Martínez JM, Jiménez-Alfaro B, Barquín J, Ondiviela B, Recio M, Silió A, Juanes JA (2017). Modelling the area of occupancy of habitat types with remote sensing. Methods in Ecology and Evolution. 9: 580-593. *Download*

Artime, I. M., Fernández, I. V., Marcos, J. B., & García, E. B. (2012). Desarrollo de un plan de gestión sostenible. Nuevos retos de la ganadería extensiva: un agente de conservación en peligro de extinción, 260.

Busqué, J. (2015). Science into practice: tools to manage mountain livestock systems and grazing commons of cantabria within the common agricultural policy. Revista Pastos, 44 (1), 6-42.

Busque, J., & Herrero, M. (2009). Atributos funcionales de las plantas forrajeras y su implicación en el manejo de pasturas. Pasturas Tropicales, 23-46.

Busqué, J., Fernández, N., & Fernández, B. (2006). A decision support tool to design rangeland sustainable grazing systems. Grassland Productivity.

Busqué, J., Mora, M. J., Bedia, J., & Fernández, B. (2007). Sobrepastoreo y degradación de pastos de puerto: quién, cuándo, dónde. Respuestas del modelo de simulación "Puerto". Actas de la XLVI Reunión Científica de Sociedad Española para el Estudio de los Pastos, 32, 37.

Chiboub, O., Kallel, A., Frison, P. L., & Lopes, M. (2018, November). Monitoring of Grasslands Management Practices Using Interferometric Products Sentinel-1. In Conference of the Arabian Journal of Geosciences (pp. 239-242). Springer, Cham.

De Vroey, M., Radoux, J., & Defourny, P. (2021). Grassland mowing detection using sentinel-1 time series: potential and limitations. Remote Sensing, 13(3), 348.

Franke, J., Keuck, V., & Siegert, F. (2012). Assessment of grassland use intensity by remote sensing to support conservation schemes. Journal for Nature Conservation, 20(3), 125-134.

Giménez, M. G., de Jong, R., Della Peruta, R., Keller, A., & Schaepman, M. E. (2017). Determination of grassland use intensity based on multi-temporal remote sensing data and ecological indicators. Remote sensing of environment, 198, 126-139.

Griffiths, P., Nendel, C., Pickert, J., & Hostert, P. (2020). Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series. Remote Sensing of Environment, 238, 111124.

Lopes, M., Fauvel, M., Girard, S., & Sheeren, D. (2017). Object-based classification of grasslands from high resolution satellite image time series using Gaussian mean map kernels. Remote Sensing, 9(7), 688.

Reinermann, S., Asam, S., & Kuenzer, C. (2020). Remote sensing of grassland production and management—A review. Remote Sensing, 12(12), 1949.

Requena, M. G. M. (2014). Mejora y aplicación del modelo de simulación "puerto" en pastos comunales de Cantabria. Proyecto Fin de Carrera. Ingeniería de Montes. Universidad de Santiago de Compostela. Available at: www.cifacantabria.org.

Sibanda, M., Mutanga, O., & Rouget, M. (2016). Discriminating rangeland management practices using simulated hyspIRI, landsat 8 OLI, sentinel 2 MSI, and VENµs spectral data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9(9), 3957-3969.

Taravat, A., Wagner, M. P., & Oppelt, N. (2019). Automatic Grassland Cutting Status Detection in the Context of Spatiotemporal Sentinel-1 Imagery Analysis and Artificial Neural Networks. Remote Sensing, 11(6), 711.

Torres, A. M., Balbi, S., & Villa, F. (2021). Scientific modelling can be accessible, interoperable and user friendly: An example for pasture and livestock modelling. bioRxiv.

Contributions of CLMS products to the Spanish LULUCF report

For this input assessment conducted by ETC-ULS-UMA, the baseline information provided by Spain was taken from the NIR 2019 in relation to definitions and stratification, where stratification is understood as the practice of dividing the key category into certain geographical strata and applying the respective stratum-specific carbon stock values and/or emission factors to calculate GHG removals/emissions. For CLMS products, efforts focused on the Corine land cover, the HRL and the CLC+ Backbone (based on information accessible in the CLMS technical library).

<u>Forest lands (FL)</u>

There are three Corine classes that can fit into the forest land category:

- 3.1.X Forest
- 3.2.4 Transitional woodland/shrub
- 3.3.4 Burnt area

HRLs, forest and small woody feature, can contribute to define these areas.

However, there are a number of limitations, both with respect to the main forest land category and to stratification.

The most important problem within the main category is a lack of equivalence between CLC and the cartography currently used in Spain, as the Minimum Mapping Unit (MMU) is different in CLC (25 ha) and in the Spanish cartography (1 ha). Another drawback in relation to equivalence arises from the use of different minimum crown cover to define forest land (30% in CLC, 50% in CLC+ Backbone and 20% according to the Spanish definition of forest land). However, these problems can be solved, in the first case by redefining the forest mask using the HRL or in the future with CLC+ Backbone, while in the case of the fraction of land cover it can be partially solved with the information on tree cover density from the Copernicus HRL, as expected in the CLC-core.

The other problem in relation to the main category is associated with the class "3.3.4 Burnt area", which is not exclusive to forest land, so that other semi-natural lands can also be included.

On the other hand, related to the limitations of the stratification, the CLMS products do not provide information on tree species or soil type.

In relation to the gaps defined for the NIR, the following CLMS products can help to cover some of the needs of the SEI:

- CLC:
    - Identify forest types (coniferous/broadleaved/mixed), without going as far as identifying tree species.
    - Detects the occurrence of reforestation, as well as monitoring the success of reforestation.

- HRL
    - Identify forest types (coniferous/broadleaved), without identifying tree species.
    - Due to its spatial resolution (10m) and update frequency (three years), it is better for identifying reforested areas and monitoring the follow-up of success in terms of density.

<u>Croplands ((CL)</u>

There are four Corine classes that can fit into the cropland category:

- 2.1.X: Arable land
- 2.2.3: Olive groves
- 2.2.1: Vineyards
- 2.2.2: Fruit tree and berry plantations

The upcoming HRLs, crop type and Phenology and productivity can support the definition of these areas.

There are several limitations, all of them in this case associated with stratification, starting with fires on cropland, which are not mapped, and furthermore it is not possible to distinguish between controlled and wild fires. On the other hand, class 2.1.X does not distinguish between fallow areas and areas currently in use, while the problem with class 2.2.2 arises both from the non-inclusion of carob and umbrella pines and the impossibility of separating citrus trees from other types of fruit tree plantations within this class. The limitation with class 2.2.3 is that it excludes wild olive trees (Oleaster spp.) as part of sclerophyllous vegetation zones, as well as abandoned olive groves. Finally, these classes cannot distinguish between cultivated land and soil types.

In relation to the gaps defined for the NIR, the following CLMS products can help to cover some of the needs of the SEI:

- CLC:
    - CLC classes include non-irrigated arable land, permanently irrigated land and rice fields, allowing differentiation between irrigated and non-irrigated land and the change between these two categories over time (2000-2018). However, CLC agricultural area often differs from other spatially explicit datasets, such as national LPIS (Parcel Identification System) or cadastres, due to the large MMU of Corine Land Cover, which often underestimates agricultural area.

Grasslands (GL)

There are three Corine classes that can fit into the grassland category:

- 3.2.3: Sclerophyllous Vegetation
- 3.2.2: Moors and heathland
- 2.3.1: Pasture, meadows and other permanent grassland under agricultural use

However, CLC does not distinguish between grassland with and without woody vegetation cover, although grassland and the phenology and productivity of HRLs can help to define these areas.

The main limitations, all related to stratification, are firstly the impossibility to distinguish between grassland with or without woody vegetation using CLC and the lack of information regarding controlled fires within the CLMS products, although it is partially covered by the category "3.3.4 Burnt areas", although this class also includes forest land and semi-natural land. Finally, there would also be a limitation in determining soil type, as the CLMS products do not provide this information for grassland or any of the LULUCF categories.

In relation to the gaps defined for the NIR, the following CLMS products can help to cover some of the needs of the SEI:

- CLC:
    - Separation into three CLC classes: sclerophyll vegetation (shrubby sclerophyll vegetation in a climax stage of development, including maquis, scrub and garrigue), moorland and heathland (Vegetation with low and closed canopy, dominated by shrubs, bushes, dwarf shrubs (heather, brambles, broom, gorse, gorse, laburnum, etc.) and herbaceous plants, which form a climax stage of development) and grasslands, meadows and other permanent pastures for agricultural use. These CLC classes can help to improve grassland mapping. However, one of the limitations of this product is that the class "3.2.4 transition woodland/shrubland" could interfere with the grassland LULUCF category, mainly for shrubland with sparse woodland.
    - In relation to the discrimination of grassland by utilisation, this product includes classes 2.3.1 (Pastures, meadows and other permanent grasslands under agricultural use) and 3.2.1 (Natural grassland) that indicate the intensity of use to a very low level of detail.
- HRL
    - HRL forest provides information on tree cover density, an element that can help in improving grassland mapping. HRLs, being thematic layers, need to be combined to obtain the full gap information, which is a limitation.
    - HRL grasslands includes includes elements of grasslands on land without use, natural, semi-natural, agricultural/managed grass-covered surfaces and grasslands with scattered trees and shrubs covering a maximum 10%. This, as in the previous case, helps in the improvement of grassland mapping. However, HRL grasslands are a Boolean map (grassland / non-grassland). Stratification by grassland types (pasture, shrubland) is not feasible. HRLs, being thematic layers, have to be combined to get the full information on this gap, which is another limitation.
    - Regarding the discrimination of grasslands by management type, the HRL grasslands includes the expert product "Ploughing indicator" (PLOUGH) which maps from 1 to 6 the number of years since the last ploughing indication at a pixel size of 10m (2018) / 20m (2015). Most of the areas that were ploughed during that 6 year period were cleared from grassland. This is an additional dataset for advanced users. However, it only serves to differentiate one type of management, without having information on irrigation, fertilisation or species improvement.

Wetlands (WL)

There are four Corine classes that may fit into the wetland category:

- 4.x Wetlands

- 5.1 Inland waters
- 5.2.1 Coastal lagoons
- 5.2.2 Estuaries

HRL Water and Wetness and the Local Coastal product can help to define these areas.

One of the limitations, as in the previous cases, is that the CLMS products do not provide information on soil type. Additionally, CLMS products do not provide information on wetland management, which is a limitation for LULUCF reporting, as only managed wetlands have to be reported.

In relation to the gaps defined for the NIR, the following CLMS products may help to cover some of the needs of the SEI (see Tables 3.20, 3.21 and 3.22):

- Local CLMS products
    - The products "Coastal zones", "Natura 2000" and "Riparian zones" allow additional stratification of wetlands and discrimination between wetlands to be covered. However, these products have their limitations, as in the case of "Coastal zones" they only cover a buffer of 10 km landward from the coast. The "Natura 2000" product only covers Natura 2000 sites and the "Riparian zones" product only covers a variable buffer of selected rivers (Strahler levels 2-9 derived from EU-Hydro).

Table 3.20. Definition of wetland classes according to the local CLMS product "Coastal Zone".

7 Wetland	7.1 Inland wetlands	7.1.1 Inland marshes	
		7.1.2 Peat bogs	7.1.2.1 Exploited peat bogs
			7.1.2.2 Unexploited peat bogs
	7.2 Coastal wetlands	7.2.1 Salt marshes	
		7.2.2 Salines	
		7.2.3 Intertidal flats	
8 Water	8.1 Water courses	8.1.1 Natural & semi-natural water courses	
		8.1.2 Highly modified water courses and canals	
		8.1.3 Seasonally connected water courses (oxbows)	
	8.2 Lakes and reservoirs	8.2.1 Natural lakes	
		8.2.2 Reservoirs	
		8.2.3 Aquaculture ponds	
		8.2.4 Standing water bodies of extractive industrial sites	
	8.3 Transitional waters	8.3.1 Lagoons	
		8.3.2 Estuaries	
		8.3.3 Marine inlets and fjords	

Table 3.21. Definition of wetland classes according to the local CLMS product "Natura 2000".

Level 1	Level 2	Level 3
7 Wetland	7.1 Inland marshes	
	7.2 Peat bogs	7.2.1 Exploited peat bog
		7.2.2 Unexploited peat bog
8 Lagoons, coastal wetlands and estuaries	8.1 Coastal wetlands	8.1.1 Coastal salt marshes
		8.1.2 Salines
		8.1.3 Intertidal flats
	8.2 Coastal waters	8.2.1 Coastal lagoons
		8.2.2 Estuaries
9 Rivers and lakes	9.1 Water courses	9.1.1 Interconnected water courses
		9.1.2 Highly modified water courses and canals
		9.1.3 Separated water bodies belonging to the river system
	9.2 Lakes and reservoirs	9.2.1 Natural water bodies
		9.2.2 Artificial standing water bodies
		9.2.3 Intensively managed fish ponds
		9.2.4 Standing water bodies of extractive industrial sites

Table 3.22. Definition of wetland classes according to the local CLMS product "Riparian Zones".

Level 1	Level 2	Level 3	Level 4
6 Sparsely vegetated land	6.2 Beaches, dunes, sands	6.2.1 Beaches and dunes	
		6.2.2 River banks	
	6.3 Bare rocks, burnt areas, glaciers and perpetual snow	6.3.1 Bare rocks and rock debris	
		6.3.2 Burnt areas (except burnt forest)	
		6.3.3 Glaciers and perpetual snow	
7 Wetland	7.1 Inland marshes		
	7.2 Peat bogs	7.2.1 Exploited peat bog	
		7.2.2 Unexploited peat bog	
8 Lagoons, coastal wetlands and estuaries	8.1 Coastal wetlands	8.1.1 Coastal salt marshes	
		8.1.2 Salines	
		8.1.3 Intertidal flats	
	8.2 Coastal waters	8.2.1 Coastal lagoons	
		8.2.2 Estuaries	
9 Rivers and lakes	9.1 Water courses	9.1.1 Interconnected water courses	
		9.1.2 Highly modified water courses and canals	
		9.1.3 Separated water bodies belonging to the river system	
	9.2 Lakes and reservoirs	9.2.1 Natural water bodies	
		9.2.2 Artificial standing water bodies	
		9.2.3 Intensively managed fish ponds	
		9.2.4 Standing water bodies of extractive industrial sites	
10 Sea and ocean			

- CLC:
    - CLC includes wetland use-related classes for the whole EEA30 territory, which helps to cover the additional stratification of wetlands, including a class for salt marshes and a class for salines. However, this product has the limitation of having a minimum mapping unit of 25 ha.
    - For discrimination between wetland types, this product includes the following categories: inland marshes, peatbogs, salt marshes, salines, intertidal flats, watercourses, water bodies, coastal lagoons and estuaries. As above, the limitations are related to the minimum mapping unit. In addition, information on wetland use is limited for some wetland types.
- HRL
    - HRL Water and Wetness provides information on permanent water, temporary water, permanent wetness and temporary wetness, which provides information on permanent/temporary water surface changes from 2012 to 2018. This information would help to fill the water body level variation gap, although information on sedimentation rates

and water residence times is not covered. This Water and Wetness layer contains defined classes of permanent water, temporary water, permanent wet, temporary wet and dry areas, derived from water and wetness occurrences in the period 2012-2018. Depending on the data needs in terms of temporality (annual biannual, etc.), this may become a limitation.

Settlements (SL)

There are eleven Corine classes that can fit into the category of settlements:

- 111: Continuous urban fabric.
- 112: Discontinuous urban fabric.
- 121: Industrial or commercial units and public facilities.
- 122: Road and rail networks and associated land.
- 123: Port areas.
- 124: Airports installations.
- 131: Mineral extraction sites.
- 132: Public, industrial or mine dump sites.
- 133: Construction sites.
- 141: Green urban areas.
- 142: Sport and leisure facilities.

HRL imperviousness and Local product Urban Atlas can contribute to define these areas.

4. Implications of Copernicus products for reporting

For estimating the areas to use additional wetland specific data sources, such as statistics from peat associations, water body databases, Natura 2000 maps. All countries that report emissions/removals of this category use some information that is being used from aerial photos, CLC and satellite images. Copernicus data of Sentinel 1 could be used for open water areas at grasslands and wetlands. These areas change during the year, so additional information data could be implemented in the reporting system at least once in 3 months.

Finland - Lessons learned

In this study foreseen potential of Copernicus products in LULUCF reporting in Finland were tested by using the CORINE Land Cover datasets from 2000, 2006, 2012 and 2018. In Finnish case high resolution CORINE data of 20 m pixel size or 25 m for older layers were available. MMU in change layers were 0.5 or 1 ha after filtering noise. We tested how well this data fit with existing field data collected by national forest inventory project (NFI), which data is used in Finnish GHG inventory. Regarding to different land categories, it was noted that CORINE takes into account land cover and partly land use and that NFI is more based on land use. Difference between land use and land cover was clearly observed in forest land. Separating forest areas from settlements was a challenging task. CORINE procedure to create settlements layer is more fixed, in NFI areas next to the build-up areas are included in settlements category if the land is not available for forest management practices - there is case by case interpretation in the field. As NFI land classification takes in the account property borders and land use, larger areas with tree cover are included in settlements than with CORINE land cover.

When comparing CORINE layers to NFI it was also noted that there were large differences in peatland forest areas. This and differences in total area of forest land should be considered because the area data need to be linked to appropriate carbon stocks and emission factors. When area estimates and carbon stock changes are estimated based on different data sources, there is a risk to break this linkage. Regarding to croplands and especially grasslands, there were large differences between NFI and CORINE.

CORINE data were available between six years interval and starting from 2000. Therefore, data were extrapolated to 1990 and real areas of land use changes prior to 2000 were not available from CORINE. Exact land use change year could not be determined, instead six years averages were only available. In Finland, land use change cases are often small by area, less than 0.5 ha. Observing also these is a challenge. Some difficulties were faced as classification and resolution between CORINE datasets were changed and due missing soil information. We merged soil database (MMU 6.25 ha) to CORINE in case of missing information on mineral and organic soils. Estimates on uncertainties for land use categories are needed.

Ireland - Lessons learned

Estimating historic land cover to support gap filling and validation measures developed by the Irish Inventory Agency (e.g., the EPA) was accomplished using a combination of Sentinel 2 satellite data and Landsat 5, Landsat

7 and Landsat 8 data. The software used to accomplish this was Google Earth Engine. The main benefit to using Google Earth Engine (GEE) was that it was a centralised location for both Sentinel and Landsat imagery. However, considerable computation power was required when utilising various image classification techniques (e.g., Object based Image Analysis (OBI)) and this limited the image classification techniques to pixel-based methods. In addition, applying OBI required considerable expertise and this was not well documented in the GEE technical guidance. Apart from this, GEE was considered an appropriate tool given the automation of many aspects of image classification techniques (e.g., creating training datasets, validating the classification using confusion matrices and many other reasons).

The main challenges identified in this case study relate to access to appropriate training data. To adequately establish land cover, point data is required to establish the correct land cover classification. Visual inspections and the use of CORINE have adequately created a reliable dataset however in the agricultural sector, considerable variation was observed when classifying grasslands and croplands. Many factors can influence the classification process from the time of year that the image was taken (e.g., the phenological stage of the crop in question or bare soil from tilled land) to land cover change between years.

One way to overcome this, is to use images at similar stages of the year (e.g., only images taken during the summer period) however cloud cover can make image acquisition challenging and this study used median pixel values from a combination of images taken over the year in question. To overcome inter-annual land cover change, access to datasets like Land Parcel Information System (LPIS) would enable a greater degree of confidence and allow greater validation. However, despite not having LPIS data, we developed a satisfactory validation process using the Land Use and Coverage Area frame Survey (LUCAS) data. This was compared to our image classifications and an estimation of accuracy was established. However, using LUCAS data was time consuming as coordinates (point data) were not always aligned with land cover and needed to be manually checked and adjusted (e.g., grasslands coordinates might be located on a hedgerow or roadside). This limited the use of LUCAS data to using small sample sizes however it remains a highly valuable tool for validation of image classifications.

Poland - lessons learned

Our study showed that the largest discrepancy in land cover changes between 2012-2015, 2015-2018 periods were observed in areas assigned to land converted to grasslands, croplands, wetlands, and settlements, where the change area between 2012-2015 was much smaller than between 2015-2018. The results of this study revealed that many land cover changes derived based on the integration of HRLs are not real changes. The main challenge identified in case of working with the Copernicus HRLs was the quality of the HRLs. The quality and spatial accuracy of HRL products varied between years 2012-2015-2018. The quality of individual layers has a strong impact on the reliability of the land cover changes. In many areas the verification of the land cover changes, derived from the comparison of the HRLs, revealed disagreement. The quality and accuracy of HRLs must be considered and further improvement of the systematic omission and commission errors in HRLs must be addressed. It needs to be stressed that the improvement of HRLs quality has been observed over time.

POLAND'S NATIONAL INVENTORY REPORT 2020 includes the vector layers for organic soils taken from *Spatial Information System on Wetlands in Poland* identified by Corine Land Cover maps for: 1990, 2000, 2006, 2012 and 2018.

Czech Republic - lessons learned

Our study showed that Copernicus Sentinel-2 data brings new perspectives in the LULUCF reporting. It is supported by a development of cloud-based technologies, e.g., Google Earth Engine (GEE). The pixel-based classification techniques using Random Forest classifier seems to be useful and the most accurate within the machine learning methods. GEE allows classification of a time high density dataset of Sentinel-2 data and automation of image pre-processing and classification techniques. Random Forest classifier needs an input of appropriate training data. The correct land cover classification is influenced by this aspect. From the point of view of LULUCF classes, the agricultural land classified into grasslands and croplands needs relevant training points and the most suitable date of Sentinel-2 image. The class of Other seems to be problematic in Czechia due to a high heterogeneity and small area. A positive of using Sentinel-2 data and cloud-based technologies is an option of annual classification of a large-scale area.

Based on our results achieved and methods used, the implementation of Sentinel-2 data classified by machine learning methods seems to be perspective, however a combination with traditionally used data (Cadastral data) is still needed for requested LULUCF reporting parameters, e.g., long period of reporting or Minimal mapping unit. Some classes classified based on Sentinel-2 data refers weak accurate results, e.g. Other land.

Concerning implementation of Copernicus data, it is necessary to collect and maintain LULUCF data in a spatial data system (GIS). If it is ensured it could be possible to combine used cadastral data with Copernicus data and services as well as with other potential data, e.g. LPIS, Land cover open data, Open street map data.

Spain - lessons learned

On this study and knowing the gaps that the LULUCF reporting currently presents in Spain, the capabilities, and opportunities of the CLMS products for monitoring and reporting LULUCF regulation, including Corine land cover classes and HRLs products, were evaluated. With this in-depth revision and the contributions of the Advisory Research Group, made up of all the entities collaborating in finding remote sensing-based solutions (Copernicus and other sources) to gaps, it was revealed that the tools provided by Copernicus could cover an important part of the gaps that MITERD was currently unable to report due to the lack of a data source, thus demonstrating the usefulness of these tools for LULUCF reporting. Moreover, by adding the contributions of the Advisory Research Group through tools related to Copernicus or other remote sensing methods, it was possible to cover the main information gaps that MITERD had defined.

One of the limitations of CLMS tools is that they do not cover the constraints related to management information. In addition, they also have certain limitations in terms of the gaps they fill, as in the case of better identification of wetland uses through the different CLMS local products, in which case they do help to improve wetland identification, but the information is limited for some types of wetlands.

Regarding the gap solutions proposed by the Advisory Research Group, these solutions were able to respond to more specific gaps through different remote sensing tools (CLMS products and other Remote Sensing techniques). However, most of the solutions proposed by the Advisory Research Group were focused on forestland gaps, covering to a lesser extent the gaps regarding the other categories (wetlands, grasslands and croplands).

In conclusion, the solutions provided by both CLMS products and other remote sensing tools help to fill the gaps in the LULUCF report, but with certain limitations, especially in the solutions to more specific gaps. These limitations may be covered to a large extent by the CLMS products that are still under development and that will help to report the relevant information in a less limited and more specific way.

Bulgaria-lessons learned

Our study shows that the application of Copernicus products will help the country to switch to a higher Tier and Approach 3: spatially explicit land-use and land use change data. The whole process for GHG emissions and removals inventory could be updated if a common approach is developed combining different methods by adding to the existing ones and spatially explicit methods that can provide transparent information on land use and land use change data.

Based on the methods applied and results achieved in our study case, a technical consistency of different data source remains the main challenge for Bulgaria. Not exact correspondence between LULUCF land use categories and land cover classes mapped with Random forest classification of Sentinel-2 , ERRORS IN MAPPING, frequent cloud cover that limited number of good Sentinel-2 imagery (method 1) and lack of rule of priority when combining HRLs, HRL corresponding to “Annual crops”, “Perennial crops”, “Pastures and meadows”, “Shrubs and grasslands”, and “Other lands” categories of the LULUCF classification in Bulgaria are gaps to be filled in near future LULUCF inventory reporting. Thus, a consolidation of available different information sources with a view to improve LULUCF inventory process should be a governmental goal to help filling in the gaps in the LULUCF report for Bulgaria.

Objectives	Bulgaria (SRT-BAS & CASTRA)	Czech Republic (CUNI)	Finland (FMI (LUKE & INRAE) & SYKE	Ireland (MU)	Poland (IGIK & CBK-PAN)	Spain (IHCANTABRIA)
Review of current reporting and monitoring LULUCF GHG	Bulgaria uses different datasets to retrieve information on land use and land use changes to get the most comprehensive information for its land representation system. Currently Bulgaria applies a combination of Approach 1 and Approach 2 , (defined in Regulation (EU) 2018/841), in elaboration of land-use change matrices. Although geo-spatial data is available in the country many technical and administrative challenges burden the use of these resources.	The methodology based on the cadastral land use information of the Czech Office for Surveying, Mapping and Cadastre (COSMC; www.cuzk.cz). The Czech land-use representation and the land-use change identification system use annually updated COSMC data, elaborated at the level of about 13 thousand individual cadastral units. The result is a system of consistent representation of land areas having the attributes of both Approach 2 and Approach 3 (IPCC 2006), permitting accounting for all land-use transitions in the annual time step. COSMC provides the annually updated areas for all land-use	The current method for monitoring land use and land-use changes is based on the network of sample plots established and measured by the National Forest Inventory (NFI). Systematic NFI sampling design covers all land uses, not only forest land. Information on land use and land-use change of each sample plot is assessed in the field and afterwards complemented with remote sensing data, digital maps, and data from the EU Land Parcel Identi-fication System	Irelands emissions/removals are currently estimated using bottom up methods collected from a mixture of state agencies. Activity data includes Land Parcel Information System (LPIS), National forest inventory ,CORINE land use data, and from the Central Statistics Office. However in 2018, the PRIME2 (classified aerial photography) that estimated full land coverage of Ireland. The introduction of this new dataset is the basis for the collaboration between MU and the national inventory agency (EPA) as described in section 3.2	The National Centre for Emissions Management (KOBIZE) is the entity directly responsible for GHG inventory. Most of the input data used in the inventory process comes from official national statistics in the statistical studies in Statistics Poland reports of Forest Management and Head Office of Geodesy and Cartography of Poland. The greenhouse gas inventory of the	The six main (level 1) land use (reporting) categories distinguished in the 2006 IPCC Guidelines are Forest Land, Cropland, Grassland, Wetlands, Settlement and Other Land. To find how the Spain defines the various land use categories in its territories and how these categories are further stratified geographically for calculating LULUCF GHG removals/emissions the MS National Inventory Reports (NIRs) submitted in 2019 and 2020 under Monitoring Mechanism Regulation (525/2013) were reviewed. The procedure used in the National Inventory Report to estimate the areas of land uses and

		categories, the Forest Management Institute (FMI) reports the recent data on forests (harvest, increment, felling, etc.) that are used in the land-use categories involving forest land.	(LPIS). The IPCC land-use category for each sample plot is derived from national land class and changes in land classes applied in the NFI with the above-mentioned supplementary data.		Land Use, Land Use Change and Forestry (LULUCF) sector covers all CO2 emissions and removals of due to gains and losses in the relevant carbon pools of the predefined six land-use categories, as well as non-CO2 emissions from biomass burning and disturbance associated with land-use conversions. All activity data, emission factors and resulting emission data are stored in a database in KOBIZE, which is constantly updated and extended to	land use changes in the period 1990-2017, is based on the exploitation of different cartographic and statistical databases available at country extent. The methodology used in the LULUCF sector follows the guidelines of the 2006 IPCC Guide, and, in part, of the 2013 Wetlands Supplement, using the emission / removal estimation algorithms, national parameters whenever possible, while in cases where such information is not available, those proposed in the 2006 IPCC Guide.

					meet the ever changing requirements for emission reporting, with respect to UNFCCC and CLTRAP as well as their protocols. Basic information on activity data regarding IPCC categories comes from Eurostat and National Statistics (GUS) databases.	
Evaluate how CLMS can improve monitoring and reporting of LULUCF	Copernicus products which could be used to supplement the spatial data available in the country could be initially identified such as: • Copernicus HRLs related to Forests – Forest	The CORINE and HRL (high resolution layers) could be used for training and accuracy assessment of the annual classification of LULUCF based on Sentinel-2 data using Random forest classifier.	One possible use for CLMS-products identified in the project was to serve as an independent monitoring system for verification of the results of the current GHGI method.	Our project used LUCAS data to validate land cover estimates. However this was time consuming where users have to extract points from a csv for your region of interest. However, this information is available at: (https://ec.europa.eu/	The HRL (high resolution layers) or the S2GLC approach for LC classification could be used for training and accuracy assessment of the annual classification of	To help evaluate the potential of CLMS products to support future LULUCF reporting, a study comparing the specifications of these products against the explicit land use category definitions and stratifications used by Spain was carried

	type and Tree cover density, Forest Additional Support Layer • Copernicus HRLs related to Grasslands - Grass Vegetation Probability Index (GRAVPI), Ploughing Indicator (PLOUGH) The application of HRLs in GHG emissions and removals estimates could be used as an independent dataset for verification of GHGI methods currently used by the responsible institutions.			eurostat/web/lucas/lucas-photo-viewer) but is not downloadable by users. This information could be made more accessible.	LULUCF based on Copernicus data. Such approach has been applied for Podlaskie NUTS2 . The emission from these areas of C02 could be compared to these reported by KOBIZE.	out. This study focussed on: 1, How Spain defines the IPCC land use categories. Stratification is understood as the practise of splitting the main category between certain geographical strata and applying respective stratum-specific C stock values and/or emissions factors to calculate GHG removals/emissions; 2, What is the current methodology used by Spain to monitor the LULUCF sector under the national inventory report (NIR); 3, What extent the products CLMS currently offered can be used for LULUCF reporting and which factors may be preventing further use of these products. Ultimately this will inform the development of a CLC+ LULUCF instance to

						support the LULUCF GHG emissions reporting as of 2023.
Comparisons of national data sets with CLMS	We compared land use areas by different information sources as reported in the national GHGI report with the areas derived by the proposed satellite-based approaches.	Our project will compare land use statistics from cadastral register with remote sensing-based classification of Sentinel-2 data. This result will be compared to CORINE and other relevant CLM datasets (e.g., high resolution layers)	We compared the national data used in the current GHGI (NFI data and additional national spatial data) to different CLMS products. In the comparison, areas of IPCC land-use categories were calculated using the different datasets and then compared. GHG emissions were also calculated using land cover and land cover change areas according to the CLMS data (Finnish national high resolution CLC data time series).	Our project will compare land use statistics obtained from Sentinel and Landsat data and compare this to CORINE and other relevant CLM datasets (e.g., high resolution layers)	The submitted project compares land use statistics which prevented KOBIZE from cadastral, National Statistics with the LULUC based on HRL and S2GLC data. According to the IPCC category National Land Identification System covers the following categories: 4.A Forest land forest 4.B Cropland arable land, orchards 4.C Grassland permanent meadows and pastures; woody and bushy land	We developed an *in-deph* revision of all capabilities and opportunities of CLMS products for monitoring and reporting LULUCF regulation. Deadline: end of May. We will disentangle how available data and methods, both from CLMS and time series of RS data (Sentinel and others), can fill the gap of data needs of Spanish end users for LULUCF.

					4.D Wetland 4.E Settlements agricultural build-up areas; build-up and urbanized areas 4.F Other land These land use categories have been compared	
Complimentary and/or Alternative methods to CLMS identified	Classification of Landsat and Sentinel-1/2 imagery was used to derive time series of land cover maps from 2012 to 2018. However, the main issues identified were: i) land cover is derived by the classification while land use is needed; ii) different spatial resolution between Landsat and Sentinel; iii) insufficient accuracy for some		We tested several top-down approaches using atmospheric inverse model data both for CO2 and CH4 fluxes, or satellite observations of photosynthesis (solar-induced fluorescence, SIF). While these complementary top-down solutions are not yet sufficiently mature to be directly utilised for national reporting,	We used LANDSAT data to estimate landcover for pre-Sentinel years. However some issues we encountered was the lack of suitable images due to infrequent satellite passes for pre-Landsat8. In addition, 30m resolution was found to very coarse for image classification.	S2GLC classification approach provides automatic solution for LC/LU classification on Pan-European scale using Sentnel-2 images. S2GLC approach could be used to generate yearly updated maps. Although the proposed solution achieved high	As stated before, Spain is currently working on the production of a consistent time series on LULUCF maps to comply with the EU Regulation 2018/841. This new generation of maps will correspond to Approach 3 of IPCC 2006 guidance. According to the information provided by the Ministry in the NIR 2020, these new maps will be implemented in the next NIR. As a result of the survey for identifying

	classes. Additional Earth observation sources of information that could be used • Hansen Global Forest Change v1.6 (2000-2018) • European Forest Fire Information System – EFFIS • Open Foris Collect Earth		developing methods for data analysis and data fusion are considered important preparation for the near future. With the planned development of the Copernicus CO_2 Monitoring and Verification Support System, atmospheric observations will become essential data sources also for LULUCF emission monitoring.		accuracy (89%), it could be further improved by inclusion of national datasets. Flexibility regarding processing rules is an additional advantage.	Copernicus-based solutions in Spain, a number of methods and approaches have been identified for mapping each of the selected gaps by a number of efficient, Copernicus-based solutions that will allow advancing the IPCC Approach and Tier currently achieved by the Spanish Inventory.
Recommendations for development of new Copernicus products (CLC+ LULUCF;HRL)	Some sub-categories of land use adopted in the GHGI of Bulgaria are difficult to map with the current CLC and HRL. For example, future development of HRL for annual crops, perennial crops, and scrublands would		Data on the following topics were identified to be needed: -Observations also on small or rare occurrences in land use and land use change -A more thorough wetland data separated for soil types and/or wetland types on	We feel that datasets like LUCAS can be expanded via increased spatial and temporal resolution and the accessibility of this data be improved. Validated classification techniques will provide increased confidence and demonstrate the effectiveness and		The provided solutions are aligned with Copernicus datasets and methods, allowing the harmonization of success stories, developed at any spatial and temporal resolution, but able to be upscaled to MS and pan EU levels by capitalizing existing knowledge and IT capabilities (e.g.

	be very convenient.		soil moisture and inundation (current HRL WaW doesn’t provide needed precision). -Data on biomass for other land uses than forest land like grasslands, wetlands and settlements	comparability to other methods.		Copernicus DIAS or any other cloud/operational system).
Others						
Remarks						

5. References

Crippa, M., Oreggioni, G., Guizzardi, D., Muntean, M., Schaaf, E., Lo Vullo, E., Solazzo, E., Monforti-Ferrario, F., Olivier, J.G.J., Vignati, E., Fossil CO2 and GHG emissions of all world countries - 2019 Report, EUR 29849 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-11100-9, doi:10.2760/687800, JRC117610.

Denier van der Gon, H. A. C., Kuenen, J. J. P., Janssens-Maenhout, G., Döring, U., Jonkers, S., and Visschedijk, A.: TNO_CAMS high resolution European emission inventory 2000–2014 for anthropogenic CO_2 and future years following two different pathways, Earth Syst. Sci. Data Discuss. [preprint], https://doi.org/10.5194/essd-2017-124, in review, 2017.

Gislason, P., & Benediktsson, J. & Sveinsson, J. (2006). Random Forests for land cover classification. Pattern Recognition Letters. 27. 294-300. 10.1016/j.patrec.2005.08.011.

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

Klemas, V. (2013). Remote Sensing of Coastal Wetland Biomass: An Overview. Journal of Coastal Research, 290, 1016–1028. Retrieved from http://www.bioone.org/doi/abs/10.2112/JCOASTRES-D-12-00237.1

Kleinen, T., Mikolajewicz, U., Brovkin, V., 2020. Terrestrial methane emissions from the Last Glacial Maximum to the preindustrial period. Clim. Past 16, 575–595. https://doi.org/10.5194/cp-16-575-2020.

Kuenen, J. J. P., Visschedijk, A. J. H., Jozwicka, M., and Denier van der Gon, H. A. C.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, Atmos. Chem. Phys., 14, 10963–10976, https://doi.org/10.5194/acp-14-10963-2014, 2014.

Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T.V., & Dech, S. (2011). Remote Sensing of Mangrove Ecosystems : A Review.

Lewiński, S., Nowakowski, A., Malinowski, R., Rybicki, M., Kukawska, E., Krupiński, M., 2017. Aggregation of Sentinel-2 time series classifications as a solution for multitemporal analysis, in: Proc. SPIE 10427, Image and Signal Processing for Remote Sensing XXIII, 104270B (4 October 2017).

Malinowski, R., Lewiński, S., Rybicki, M., Gromny, E., Jenerowicz, M., Krupiński, M., Nowakowski, A., Wojtkowski, C., Krupiński, M., Krätzschmar, E., Schauer, P., 2020. Automated production of a land cover/use map of Europe based on sentinel-2 imagery. Remote Sens. 12, 1–27.

McLeod, E., Chmura, G. L., Bouillon, S., Salm, R., Björk, M., Duarte, C. M., Lovelock, C. E., Schlesinger, W. H., & Silliman, B.R. (2011). A blueprint for blue carbon: Toward an improved understanding of the role of vegetated coastal habitats in sequestering CO2. Frontiers in Ecology and the Environment, 9(10), 552–560.

Official Statistics of Finland (OSF): Greenhouse gases [e-publication]. ISSN=1797-6065. Helsinki: Statistics Finland [referred: 23.4.2021]. Access method: http://www.stat.fi/til/khki/index_en.html

Raivonen, M., Smolander, S., Backman, L., Susiluoto, J., Aalto, T., Markkanen, T., Mäkelä, J., Rinne, J., Peltola, O., Aurela, M., Lohila, A., Tomasic, M., Li, X., Larmola, T., Juutinen, S., Tuittila, E.-S., Heimann, M., Sevanto, S., Kleinen, T., Brovkin, V., Vesala, T., 2017. HIMMELI v1.0: HelsinkI Model of MEthane buiLd-up and emIssion for peatlands. Geoscientific Model Development 10, 4665–4691. https://doi.org/10.5194/gmd-10-4665-2017

Spahni, R., Wania, R., Neef, L., van Weele, M., Pison, I., Bousquet, P., Frankenberg, C., Foster, P.N., Joos, F., Prentice, I.C., van Velthoven, P., 2011. Constraining global methane emissions and uptake by ecosystems. Biogeosciences 8, 1643–1665. https://doi.org/10.5194/bg-8-1643-2011

Stocker, B.D., Spahni, R., Joos, F., 2014. DYPTOP: a cost-efficient TOPMODEL implementation to simulate sub-grid spatio-temporal dynamics of global wetlands and peatlands. Geoscientific Model Development 7, 3089–3110. https://doi.org/10.5194/gmd-7-3089-2014

Tanneberger, F., Tegetmeyer, C., Busse, S., Barthelmes, A., and 55 others, 2017. The peatland map of Europe. Mires and Peat 1–17. https://doi.org/10.19189/MaP.2016.OMB.264

Tsuruta, A., Aalto, T., Backman, L., Hakkarainen, J., van der Laan-Luijkx, I.T., Krol, M.C., Spahni, R., Houweling, S., Laine, M., Dlugokencky, E., Gomez-Pelaez, A.J., van der Schoot, M., Langenfelds, R., Ellul, R., Arduini, J., Apadula, F., Gerbig, C., Feist, D.G., Kivi, R., Yoshida, Y., Peters, W., 2017. Global methane emission estimates for 2000–2012 from CarbonTracker Europe-CH4 v1.0. Geoscientific Model Development 10, 1261–1289. https://doi.org/10.5194/gmd-10-1261-2017.

Zhang, M., Ustin, S.L., Rejmankova, E., & Sanderson, E.W. (2019). Monitoring Pacific Coast Salt Marshes Using Remote Sensing. Ecological Applications, 7(3), 1039–1053.

Zomer, R.J., Trabucco, A., & Ustin, S.L. (2009). Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. Journal of Environmental Management, 90(7), 2170–2177. Retrieved from http://dx.doi.org/10.1016/j.jenvman.2007.06.028.