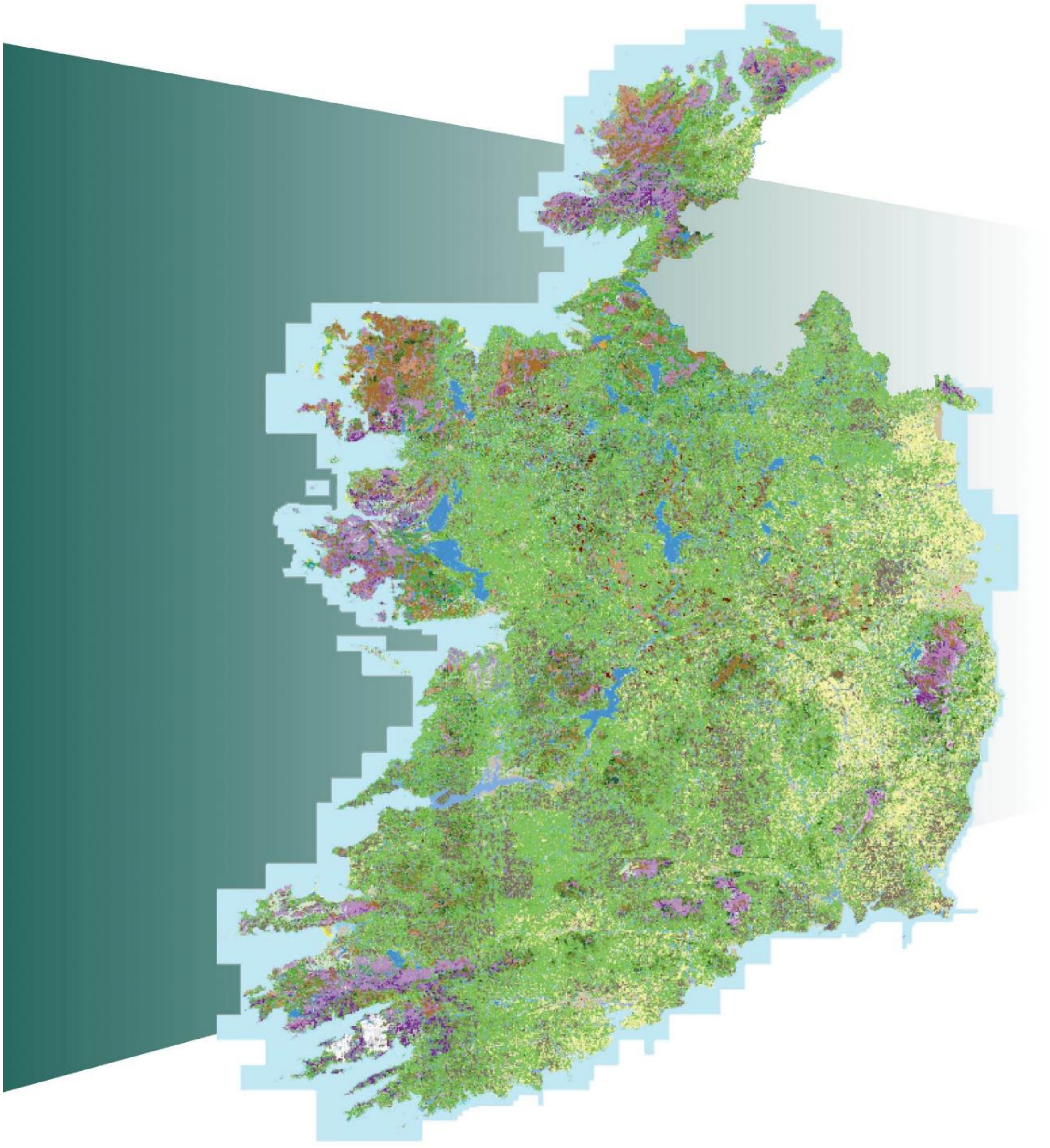
National Landcover Map of Ireland

2018

Final Report



Lead Author: Kevin Lydon (EPA)

Co Author: Gavin Smith (EPA)

Contributors:

National Mapping Division of Tailte Éireann: Rachael Kelly, Bryan Sweeney, Stephen Mullins, Joanna Chrachol, Josh O' Sullivan Hourihan and Geraldine Kavanagh.

CSO /EPA: Tomás Murray (CSO & EPA)

Introduction:
This project was delivered through a partnership between the Ordnance Survey Ireland (OSi) and the Environmental Protection Agency (EPA). It is noteworthy that subsequent to the completion of this project, Ordnance Survey Ireland (OSi) underwent a merger with the Property Registration Authority (PRA) and the Valuation Office to form Tailte Éireann - a new Government agency operating under the Department of Housing, Local Government and Heritage. Tailte Éireann will provide a property registration system, property valuation service, and the national mapping and surveying infrastructure for the State. From March 1, 2023, Ordnance Survey Ireland will be known as Tailte Éireann's National Mapping Division. We refer to Ordnance Survey Ireland (OSi) throughout this document in order to provide a more accurate representation of the actual project timeline, except where it was necessary to update contact information.

Acknowledgements:

The project team consisted of Kevin Lydon and Gavin Smith from the EPA, and Rachael Kelly, Geraldine Kavanagh, Bryan Sweeney, Guillermo Castro Camba, Joanna Chrachol, Stephen Mullins, Linda Crowley, Ruth Browne, Michael Moriarty, Monika Pienkos, Aidan Burke, Paul Kane, and Josh O'Sullivan Hourihan from the OSi. The project team were also directly supported by Simon Barron and Fionnuala O'Neill from BEC on ecological matters and Tomás Murray, a seconded CSO statistician to the EPA, on the validation and accuracy assessment. The project team would also like to acknowledge the support provided by the internal validators, ICT teams, and senior managers in both the EPA and OSi.

This project is the culmination of over a decade's work by environmental and geo-spatial experts in Ireland. Working within the National Land Cover and Habitats Mapping (NLCHM) working group these experts helped to develop and promote the case for a high-resolution land cover map for Ireland. These and many other staff in the public and private sector have also helped in the design, production, and validation of the land cover data product. We wish to acknowledge their invaluable contribution to help deliver this critical environmental dataset for Ireland.

Our thanks to supporting staff in:

- Environmental Protection Agency
- Ordnance Survey Ireland
- National Parks & Wildlife Service
- Teagasc
- Department of Agriculture, Food, and the Marine
- Central Statistics Office
- Atlantic Technical University
- Botanical, Environmental & Conservation (BEC) Consultants Ltd
- The Heritage Council
- Geological Survey of Ireland

In particular we would like to recognise the considerable input and support of the following people during the course of this project: Fiona O'Rourke, Jonathan Derham, Eimear Cotter and Matt Crowe in EPA; Colin Bray and Peter Hallahan in OSi; Gemma Weir, Gareth John and Ciaran O'Keeffe in NPWS; Stuart Green in Teagasc; Frank Barrett, Fran Morrin and Eoin Dooley in DAFM; Tom Healy, Nova Sharkey and Sylvie Clappe in CSO; James Moran and Ruth Bennett Coady in ATU; Jim Martin in BEC; Beatrice Kelly and Michael Starrett in the Heritage Council; and Ray Scanlon and Michael Sheehy in the GSI.

Our thanks and appreciation to all organisations that made data available for the production and validation of the National Land Cover Map 2018 (NLC 2018).

The project team would also like to pay tribute to the late Julie Fossitt, from the National Parks and Wildlife Service, who's exemplary work on "A Guide to Habitats in Ireland" (2000) was the initial basis for developing the National Land Cover Classification System used in this map.

And finally, a thank you to George McHugh, retired EPA, who planted the seed all those years ago.

Acronyms

ATU	Atlantic Technical University
BEC	Botanical, Environmental & Conservation Consultants Ltd
BGRs	Biogeographical Regions
CLC	Co-ORdinated INformation on the Environment
CLC+	Co-ORdinated INformation on the Environment (Next Generation)
CLMS	Copernicus Land Monitoring Service
COPERNICUS	European Space Programme
CORINE	Co-ORdinated INformation on the Environment
CORINE+	Co-ORdinated INformation on the Environment (Next Generation)
CSO	Central Statistics Office
DAFM	Department of Agriculture, Food, and the Marine
DIPM	Derived Irish Peatland Map
DSM	Digital Surface Model
DTM	Digital Terrain Model
EAGLE	EIONET Action Group on Land Monitoring in Europe
EEA	European Environment Agency
EIONET	Environment Information and Observation Network
EO	Earth Observation
EPA	Environmental Protection Agency
ESA	European Space Agency
ESPG	European Petroleum Survey Group (Series of Geographic Projection Systems)
GHG	Greenhouse Gas Emission
GIS	Geographic Information Systems (GIS)
GSI	Geological Survey of Ireland
GUID	Globally Unique Identifier
HA	Hectare
IG	Irish National Grid (Geographic Projection System)
ITM	Irish Transverse Mercator (Geographic Projection System)
KIA Score	Kappa Index of Agreement Score (A Metric of Accuracy)
LPIS	Land Parcel and Information System
MMU	Minimum Mapping Unit
NFI	National Forest Inventory
NGO	Non-Governmental Organisations
NHA	Natural Heritage Area
NLC 2018	National Land Cover Map 2018
NLCHM	National Land Cover and Habitat Mapping Working Group
NPWS	National Parks & Wildlife Service
NSDM	Normalised Surface Difference Model
OA	Overall Accuracy (A Metric of Accuracy)
OBIA	Object-based Image Analysis
OSI	Ordnance Survey Ireland
PRIME2	Name of National Cadastral Map
QGIS	Quantum GIS - Open Source GIS Software
SAC	Special Area of Conservation
SPA	Special Protection Area
Sqm	Square Metre

Table of Contents

Executive Summary

In 2018, the Ordnance Survey of Ireland (OSi) and the Environmental Protection Agency (EPA) initiated a new national land cover mapping programme for Ireland under a partnership agreement. The concept and proposal for this programme was developed over the previous decade by members of the cross-agency National Land Cover and Habitats Mapping working group (NLCHM). In 2016 this group presented a joint memo to Government proposing a detailed national land cover map. Following various consultation with Government Departments and Agencies, work began on producing a new high-resolution National Land Cover Map for Ireland in 2018 and was completed in September 2022.

Land cover maps are Geographic Information System (GIS) datasets that map the different types biological and physical materials on the earth's surface. Type of land cover include grasslands, forest lands, peatlands, artificial areas, waterbodies, and exposed surfaces within the landscape. They provide an invaluable digital environmental record of the terrestrial environment and how it changes over time due to human activity and natural processes.

Land cover datasets are used as a core environmental data source for a range of important applications including: reporting on EU environmental laws and directives; estimating national Greenhouse Gas (GHG) emissions for Land Use, Land Use Change and Forestry (LULUCF) inventories; Appropriate and Environmental Impact Assessments (AA's and EIA's), environmental noise modelling; Illegal waste detection; catchment risk assessments for water quality; State of the Environment Reports (e.g. EPA, 2020); academic research; and habitat mapping and natural capital accounting.

To date, the only source of national-scale land cover data in Ireland has been the European Environment Agency's CORINE (Co-ORdinated INformation on the Environment) land cover data series, delivered under the Copernicus Land Monitoring Service. This data, whilst providing an important record on the environment dating back to 1990, is not suitable for use in more detailed land assessments due to its low spatial resolution of 25ha and its classification system, designed on continental European land cover, that does not give sufficient detail on grasslands, peatlands and other important land cover types of Ireland.

Recognising the need for more detailed land cover data, designed specifically for the Irish landscape, a consortium of public-sector environmental and geospatial experts formed the National Land Cover and Habitat Mapping (NLCHM) working group to develop a technical approach and business case for the development of a national land cover mapping programme for Ireland. In 2016, this group brought a successful joint memo to government proposing the development of the first national land cover map. Work on the map began in 2018 led by OSi in partnership with the EPA.

Imagery from the OSi aerial photography campaign of 2018 formed the main data source alongside satellite imagery from the European Space Agency (ESA) Sentinel 2 programme and OSi's PRIME 2 vector spatial database. A novel remote sensing-based methodology was designed to classify over 10 million features in the landscape using a combination of rule-based and machine learning classification approaches.

A new land cover classification system was developed to accurately describe the different land cover types present in Ireland. Widespread consultation with expert stakeholders within a variety of sectors in Ireland was undertaken to ensure the classification system met a wide range of user needs. The classification system, designed for use in this project, also provides a common standard and framework which can be used in other studies undertaking land cover mapping in Ireland.

The National Land Cover Map of Ireland for the year 2018 (NLC 2018) is now complete. Enquiries on how to gain access to the data should be directed through the National Mapping Division (formerly OSi) of the newly established Tailte Éireann by contacting corporatesales@tailte.ie. It gives highly detailed and validated information on the thirty-six land cover types identified within the classification system. For the first time, it provides a comprehensive, national-scale mapping of all land cover types of the country including broadleaved woodland and hedgerows, a range of peatlands, wet and dry heathland, wet and dry grasslands, artificial surfaces, exposed rock and sediments.

A rigorous external validation campaign ran in parallel to the classification programme. This involved sourcing and training up to 54 validators who were expert staff from various public, private and research organisations. These validators collectively manually assessed and validated over 22,000 stratified random validation samples.

Validation figures show that a high level of overall thematic accuracy, as F1 score, was achieved with 89% of samples being classified correctly in the Level 1 group and 79% of samples being correctly classified at the more detailed Level 2. An assessment of the spatial delineation of the land cover polygon showed that all classes had a geometric accuracy >80% with the majority >85%.

More detailed analysis is to follow on the breakdown of the validation statistics which will help inform where the classification process needs to be refined in future iterations.

NLC 2018 presents new baseline information on the status of the landscape and environment in Ireland. It is hoped that this will help inform and enable more accurate and reliable assessment, reporting, research, and monitoring of the environment, while helping Ireland to tackle the many challenges presented by the Climate and Biodiversity emergencies. It is envisaged that this will be the first in a long-term series of national land cover data, which through time will help us identify and manage the changes and pressures on Ireland's Environment.

1 Introduction

1.1 Land Cover, Land Use and Land Monitoring

Land monitoring consists of a group of activities that observe and record the lands' surface, the natural processes and human activity that occurs on it and how these factors interact to change the landscape and environment over time. The most common land monitoring activities are land cover and land use mapping but also includes habitat mapping and the production of spatial datasets on specific environmental topics such as tree cover density, soil sealing and phenology mapping.

Land Cover Mapping

Land cover is defined as the 'biological and physical (bio-physical) material present on the earth's surface at the time of observation' (Di Gregorio, 2005). This refers to the different vegetation types such as grasslands, forest lands, peatlands, and non-vegetated surfaces such as waterbodies, artificial surfaces, exposed rock and sediments that occur within the landscape.

Land cover mapping is the activity of observing, describing, and recording these surface types. It is typically done through remote sensing analysis of aerial or satellite-sourced Earth Observation (EO) imagery, often in combination with in-situ data on road networks, building outlines, forestry parcel data, etc. Land cover maps are produced as raster or vector Geographic Information Systems (GIS) datasets which delineate real-world features and categorises them to a defined classification system.

Land cover maps provide a valuable, independent record of what is present in the landscape and how this change over time through iterative change-mapping. It is an evidence-base which can be used for a wide range of applications including in the preparation of national inventories on Greenhouse Gas (GHG) emissions, fulfilling national reporting requirements on a range of EU Directives including the Noise, Nitrates, Habitats and Water Framework directives. It can be used to inform national policy decisions, local infrastructure and development plans and it provides a key data source for environmental research.

Land Use Mapping

Land use mapping is another form of land monitoring that maps and records what an area of land, building or other structure is used for by humans as part of a recreational, social, or economic activity. Common land use activities include pasture and arable agriculture, commercial and residential use of buildings and recreational activities such as field sports, hill-walking, and swimming.

Land use is similar to land cover in that it describes what is occurring at a given point in the landscape, however it's the relationship with a single location on the ground is more complex. As can be seen in Figure 1.1 below, an individual land parcel can only have a single land cover class name, but multiple land use activities can take place on or within it. A single land use activity can also occur across multiple adjoining land parcels or buildings. This coupled with the fact that land use is often not visually discernible in EO data means that land use mapping often requires the use and integration of multiple sources of inter-disciplinary information, this can be spatial / GIS data or non-spatial data such as census records, grant payment forms, traffic management data, survey data etc.

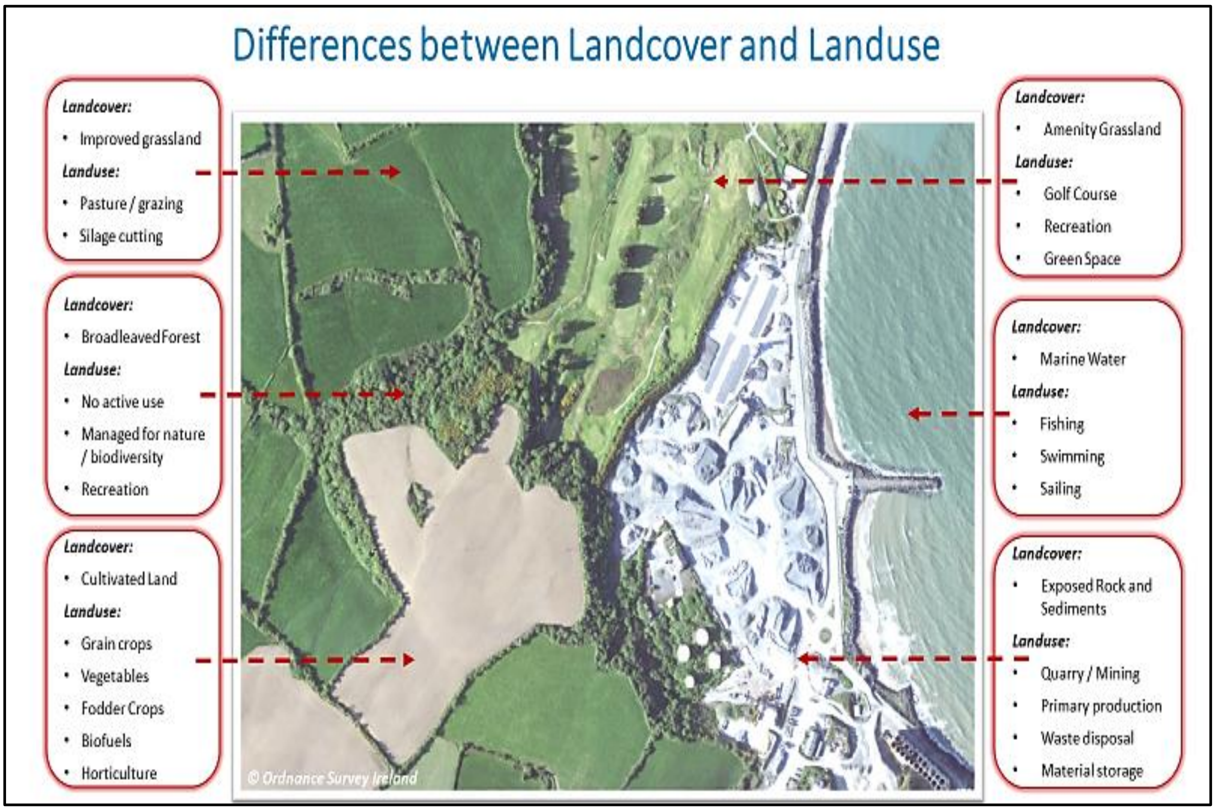


Figure 1.1: Showing different land cover types with multiple potential land use activities

1.2 Development of an Irish Land Cover Mapping Programme

National and European Land Monitoring

In the last decade, there has been marked developments in remote sensing and Earth Observation globally. In Europe, COPERNICUS - the European Commissions' flagship Earth Observation programme, has driven this through dramatically increasing the availability of EO data and services to the sector. The Sentinel satellite constellation, launched in 2014, provides freely available high resolution multispectral and radar satellite imagery from across the globe on a 4-5-day revisit schedule. The Copernicus Land Monitoring Service (CLMS), one of 5 service pillars in the Copernicus programme, provides a range of downstream land data products and services including the Pan-European high-resolution layers, the urban atlas, snow and ice cover monitoring, riparian zone mapping and the next generation CORINE product called CLC+. Copernicus also provides cloud-based Data Information Access Service (DIAS) infrastructures which allow users to access and analysis EO data products, including Sentinel imagery.

These improvements in data and services have led to a revision in how member states describe and classify the landscape, as discussed in section 3, and in the technical approaches used to mapping land cover. Overall, a more detailed level of spatial and temporal resolution is now achievable by all countries with high resolution and continuous mapping of the landscape possible.

Land Cover Mapping in Ireland

All European member states undertake some form of national-scale land monitoring, either through the Pan-European CORINE land cover mapping programme or through their own national land cover or land use mapping programmes.

To date, the only attempt to produce a national scale land cover map was the 1995 Teagasc Land Cover Map (Green, 1995). Apart from this, the only form of national-scale land cover mapping in Ireland has been through the pan - European CORINE land cover programme (Büttner et al., 2021). The first iteration of CORINE was in 1990 and several iterations have followed in 2000, 2006, 2012 and 2018. Since 2000 the EPA have produced CORINE data for the European Environment Agency.

It is widely accepted that CORINE data is a better fit for pan European assessment while not meeting the national, regional, and local assessment requirements of Irish stakeholders. This is due firstly to its low spatial resolution or minimum mapping unit (MMU) of 25 hectares (ha) (EEA, 2021), and secondly the CORINE classification system is designed to map Pan-European land cover meaning it does not adequately describe important Irish land cover classes, for example types of grassland and peatland.

This situation led to many agencies and departments undertaking their own, more detailed, local-scale and sector specific land mapping work to fulfil their own data requirements. Whilst significant public resources were being dedicated to land monitoring in Ireland overall, there was no consistent, standardised and verified data on the Irish landscape and how it was changing over time.

The National Land Cover and Habitat Mapping (NLCHM) Working Group

To address this key environmental data gap, a cross-agency working group of land cover, land use and habitat mapping experts from across the public sector was set up in 2012. A shared work programme was undertaken by the group to develop a new Irish land cover classification system, a production methodology, spatial data model and a pilot study to investigate and prove that a high quality national land cover dataset could be delivered to better serve the environmental data needs of the country.

The group brought a memo to Government in 2016 to recommend the establishment of a national land cover programme. Following this, the Ordnance Survey of Ireland (OSi) undertook to produce the dataset with the Environmental Protection Agency (EPA) as project partners. The new project commenced in late 2018 and the first National Land Cover Map of Ireland is now completed. It is based on 2018 data and is also referred to as the NLC 2018. Figure 1.2 below shows the timeline for developing the new national land cover mapping programme and the first delivery of NLC 2018.

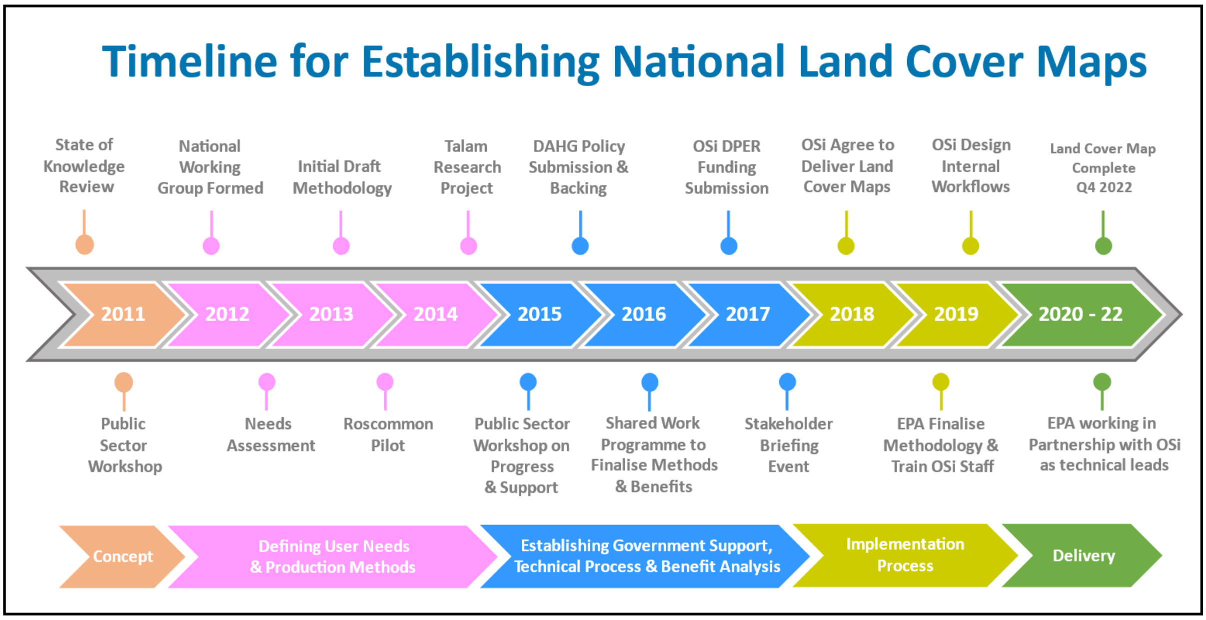


Figure 1.2: The timeline for establishing the national land cover programme in Ireland

1.3 Report Outline

Many new systems, technical solutions and standards had to be designed to successfully produce the first ever high-resolution land cover dataset for Ireland, NLC 2018. This report details the development, design, and implementation of these technical approaches within the various stages of the project. An overview of a new Land Cover Classification System for Ireland is also given. This system was a key requirement for producing a National Land Cover Map and it also provides a common framework and standardised terminology for describing the Irish landscape which other studies can utilise. The report describes how very large quantities of source data had to be collated, processed, and managed in preparation for the classification analysis. Details on the classification methodology are given, where a combination of GIS data integration and object-oriented EO image analysis are used to produce the final land cover classification.

To verify the accuracy of the data product, an unprecedented level of external validation for an Irish environmental dataset was undertaken. The report details this parallel activity and the initial validation results from it. Information on the data structure, format and how to access the data are also given.

The main stages of the project are listed below and detailed in the respective sections of the report:

1. Project setup (Section 2)
2. Design of a new Irish Land Cover Classification System (Section 3)
3. Data collation & management, software, and hardware setup (Section 4)
4. Land Cover Classification and post-processing of data (Section 5)
5. Independent external validation of the NLC 2018 data product (Section 6)

2 Project Setup and Structure

This was a large-scale and multi-faceted project, not undertaken before in Ireland, involving staff working together across different public agencies. A significant amount of time was therefore dedicated in the early stages of the project to designing and building new systems and workflows, developing solutions on technical problems, interacting with external stakeholders, testing, and refining the final data product.

2.1 Project Scope and Terms of Reference

An initial meeting of the NLCHM steering group was held to agree the scope and terms of reference of the project. OSi undertook responsibility for the production and delivery of the new land cover map for Ireland and it was decided that their high-resolution data catalogue would form the core data source for the project. The EPA, having led the development of the technical approaches to land cover mapping in the NLCHM group, joined the OSi as project partners to aid the successful delivery of the project. The remaining members of the NLCHM group participated as members of a steering group to oversee and provide expert guidance to the project, with NPWS remaining as chair of this group. The full list of organisations and departments represented on the NLCHM group are:

- National Parks & Wildlife Service (NPWS)
- Ordnance Survey Ireland (OSi)
- Environmental Protection Agency (EPA)
- Teagasc
- Department of Agriculture, Food & the Marine (DAFM)
- Central Statistics Office (CSO)
- Geological Survey Ireland (GSI)

The project scope was agreed to produce and deliver a new national land cover mapping dataset for Ireland, to the standards agreed previously by the NLCHM working group using the Ordnance Survey's data catalogue as the primary data source.

2.2 Project Start-Up and Design of Work Packages

A series of project team start-up meetings were held to determine the overall project plan, individual work packages, tasks and other issues that needed to be addressed.

An outline project plan was drawn up which set out the key work packages of the project, allowing for a significant pre-production phase where the specific tasks and details of each work package were developed out and a lead agency assigned to each task.

The main work packages were:

- Work Package 1: Project setup – see below
- Work Package 2: Design of Classification System – Chapter 3
- Work Package 3: Data pre-processing – Chapter 4.1
- Work Package 4: Land cover classification – Chapter 4.2
- Work Package 5: Data post-processing – Chapter 4.4
- Work Package 6: Data validation – Chapter 6

2.3 Project Setup

The pre-production phase involved setting up the team and project along with the following tasks:

2.3.1 Data Scoping and Collation

A key task to complete early on was to scope out and identify what Earth Observation (EO) and other GIS data was required and available to the project. This informed the level of classification detail that could be achieved, the software needed to undertake the classification and the training requirements for staff. A list of minimum data requirements was drawn up based on the pilot work undertaken in the NLCHM group and a process of identifying new data sources and requesting access to external data was undertaken.

The full OSi data catalogue was available to the project including the 'Series 2' 2018 25cm Aerial Photography collection and their 1m digital elevation models. This was a significant development from the pilot work undertaken by the NLCHM group which used Sentinel 2, 10 metre resolution imagery as the main data source. Along with this, there was a wide range of in-situ vector data made available to the project from local authorities, NGO's, and other stakeholder groups.

The data collation also involved a screening process to assess the usability of all datasets obtained. Over sixty separate datasets were made available to the project but only a small percentage were suitable for use in the project due to various reasons including incomplete or inconsistent coverage, time elapsed since the data was collected, incompatible classification systems and unreliable accuracy. A conservative approach was taken by the project in using external data sources to avoid the importation of errors or inaccuracies. Ultimately, the project had to balance the value offered by in-situ data with its reliability and the effort involved in integrating it into the processing models.

For all external data sources, data sharing agreements or memoranda of understanding were drawn up and secured between the OSi, EPA and relevant agencies. The openness to sharing data played a significant role in the success of the project.

The source data which passed the screening stage was split into two categories – Core data and secondary data.

Core Data

Core data refers to data that was used as a direct data source in the classification and production of the land cover dataset. This includes the main aerial and satellite-based EO imagery and cadastral vector datasets which were used in the classification process, including the Land Parcel Information System (LPIS) which is a spatial database of the land use of agricultural land parcels in Ireland (Zimmerman et. al 2016).

The core datasets are listed in Table 2.1 below with data descriptions included.

Secondary Data

Secondary data refers to data that was used as a reference to guide and inform the classification or validation team but not used directly in the classification process. This data includes a wide range of local-scale habitat, land use and other external data that was offered to the project. Table 2.2 shows the secondary data sources that were used in the land cover mapping project.

Dataset Name	Owner	Description	Data Type	Format	Data Size	Capture Date
Series 2 Ortho-photography	OSi	4 band colour infrared (CNIR) ortho-photography raster images captured at 25cm resolution with 15cm resolution in Dublin.	Raster	.img	5TB	2016-2019
Series 2 DTM	OSi	Photogrammetry based DTM captured at 1-meter resolution.	Raster	.img	122GB	2016-2019
Series 2 DSM	OSI	Photogrammetry based DSM captured at 1-meter resolution	Raster	.img	480GB	2016-2019
Sentinel 2 Imagery	ESA	Spring and Summer image set with bands 2, 3, 4 and 8 from the ESA Sentinel 2 sensor, captured at 10-meter resolution.	Raster	.tif	25GB	2018
PRIME 2	OSi	National cadastral vector database providing spatial boundaries for all buildings, roadways, field boundaries and water bodies.	Vector	.fgdb	19GB	Q4 2018
LPIS	DAFM	Agricultural land use dataset compiled by DAFM for the Basic Farm Payment scheme.	Vector	.fgdb	1.5GB	2018
Forest Service	DAFM	Spatial data on all grant-aided private forestry in the state. Includes information on forest species, planting date, etc.	Vector	.fgdb	< 1GB	2017
COILLTE Land Use Map	Coillte	Spatial map of all Coillte forestry and land holdings, including forest type, unplanted areas, felling and burnt areas.	Vector	.fgdb	< 1 GB	2018
Series 2 Seamlines	OSi	Vector dataset of the spatial extent and capture date for all image strips in the OSi series 2 aerial photography catalogue.	Vector	.fgdb	< 1 GB	2018
Coastal Monitoring Project	NPWS	Survey of sand dunes and other coastal habitats in Ireland between 2004-2006.	Vector	.shp (polygon)	< 1 GB	2004 - 2006
Sand Dunes Monitoring Project	NPWS	Survey of sand dune systems in Ireland	Vector	.shp (polygon)	< 1 GB	2013
Inventory of Irish Coastal Lagoons	NPWS	Series of surveys on coastal lagoons in Ireland carried out between 1996 and 2006.	Vector	.shp (polygon)	< 1 GB	2007
Saltmarsh Monitoring Project	NPWS	Survey data of over 130 saltmarsh systems around the Irish coastline.	Vector	.shp (polygon)	< 1 GB	2006-2008
National Survey of Native Woodlands	NPWS	Spatial data from a survey of over 1,200 native woodland sites in Ireland.	Vector	.shp (polygon)	< 1 GB	2003 - 2008
Ancient and long-established Woodland Inventory	NPWS	A dataset of over 480 woodland sites that have been continually covered in woodland since C. 1830.	Vector	.shp (polygon)	< 1 GB	2010

Table 2.1: Core data sources used in the classification process

Dataset Name	Owner	Description	Data Type	Format	Data Size	Capture Date
2019 Article 17 Reporting data	NPWS	Habitat mapping of Annex 1 habitats in Ireland under the Habitats Directive	Vector	.shp (polygon)	< 1 GB	2019
Site Specific Conservation Objectives (SSCO)	NPWS	Habitat maps of Natura designated sites in Ireland	Vector	.shp (polygon)	< 1 GB	2019
SAC, SPA, NHA & pNHA Boundaries	NPWS	Boundaries of all Natura designated sites in Ireland	Vector	.shp (polygon)	< 1 GB	Various
National Survey of Upland Habitats	NPWS	Habitat survey of several upland areas in Ireland	Vector	.shp (polygon)	< 1 GB	2010-2014
Irish Semi-Natural Grasslands Survey (ISGS)	NPWS	Survey of known natural and semi-natural grassland sites in Ireland	Vector	.shp (polygon)	< 1 GB	2008-2012
National Survey of Limestone Pavements	NPWS	Survey of Karst Limestone pavement areas	Vector	.shp (polygon)	< 1 GB	2013
RBMA dataset	NPWS	Habitat maps of raised bog sites	Vector	.shp (polygon)	< 1 GB	2007-2013
BnM_Fossitt_Level_2_Final300120	BNM	Habitat data for several Bord na Mona-owned peatland complexes	Vector	.shp (polygon)	< 1 GB	2018
Local Authority Habitat & Wetland maps	LA's (Various)	Local scale habitat maps obtained from Clare, Cork, Kerry, Leitrim, Wicklow, Waterford, Roscommon, and Monaghan County Councils.	Vector	.shp (polygon)	< 1 GB	Various
National Vegetation Database records	NBDC	Point data of vegetation and associated habitats	Vector	.shp (points)	< 1 GB	Various
Wetland Survey Ireland database	WSI	Database of over 2,000 wetland sites in Ireland, compiled over a 10-year period	Vector	Webservice	-	Various

Table 2.2: Secondary data sources used as reference data

2.3.2 Data Management

With the data to be used in the project identified, a plan to store and manage the large catalogue of data was agreed between the two agencies with secure servers dedicated to store and process the large datasets required for the project. An inventory of the data accessed and used in the project was kept and data sharing agreements or MOUs were signed with external agencies to secure data access.

2.3.3 Design of Working Units and Project Rollout

Considering the scale of the project in terms of geographic coverage and the quantity of data involved, it was imperative to organise the data into manageable and logical working units to provide an ordered structure and framework for the data processing, classification and validation work across the country.

The data was therefore organised into working units based on the system used by OSi to capture the main source imagery, the OSi aerial photography and the Sentinel 2 EO imagery.

Working Blocks and Seamlines

The core aerial imagery data was collected by the OSi aerial photography team in a grid of 29 working blocks across the country which can be seen in Figure 2.1 below. These blocks were then used to divide the country into manageable working units for the land cover project.

Each block was composed of a set of 15-20 image strips with a footprint of approx. 5km x 50km (see also Figure 2.1). Once captured, the raw image strips were triangulated to produce an ortho-image, edge-matched and merged to form a single raster image for the full working block. In an ideal scenario, the image strips were captured within a narrow timeframe and offer a temporally homogenous set of image data. However, due to the persistent cloud cover in Ireland, the blocks are usually composed of image strips from several flight attempts over different dates.

The line along which the image strips were edge matched together was captured as a vector dataset with the capture date for each strip attributed. As detailed in chapter 4, this ‘seamline’ dataset was a critical dataset and was used in the classification process to identify and split the image data based on the flight date of the source imagery, creating a temporally consistent classification scene.

Biogeographical Regions

The 29 working blocks were further grouped into eight regions called ‘Biogeographical Regions’ (BGRs) to organise the roll out of the project. Each BGR consisted of 3-5 working blocks, these are sub-unit blocks within the BGRs which had broadly similar land cover and topography. The BGRs were also closely aligned with the UTM grids which the footprints of the Sentinel imagery are based on. An image showing the working blocks and the eight biogeographical regions is shown in Figure 2.2 below.

Production Rollout

After initial production testing in working block 19, the Wicklow area, the production was then rolled out across the rest of the country in the following sequence:

1. South East
2. North East
3. Midlands
4. South West
5. South
6. Burren
7. West
8. North West

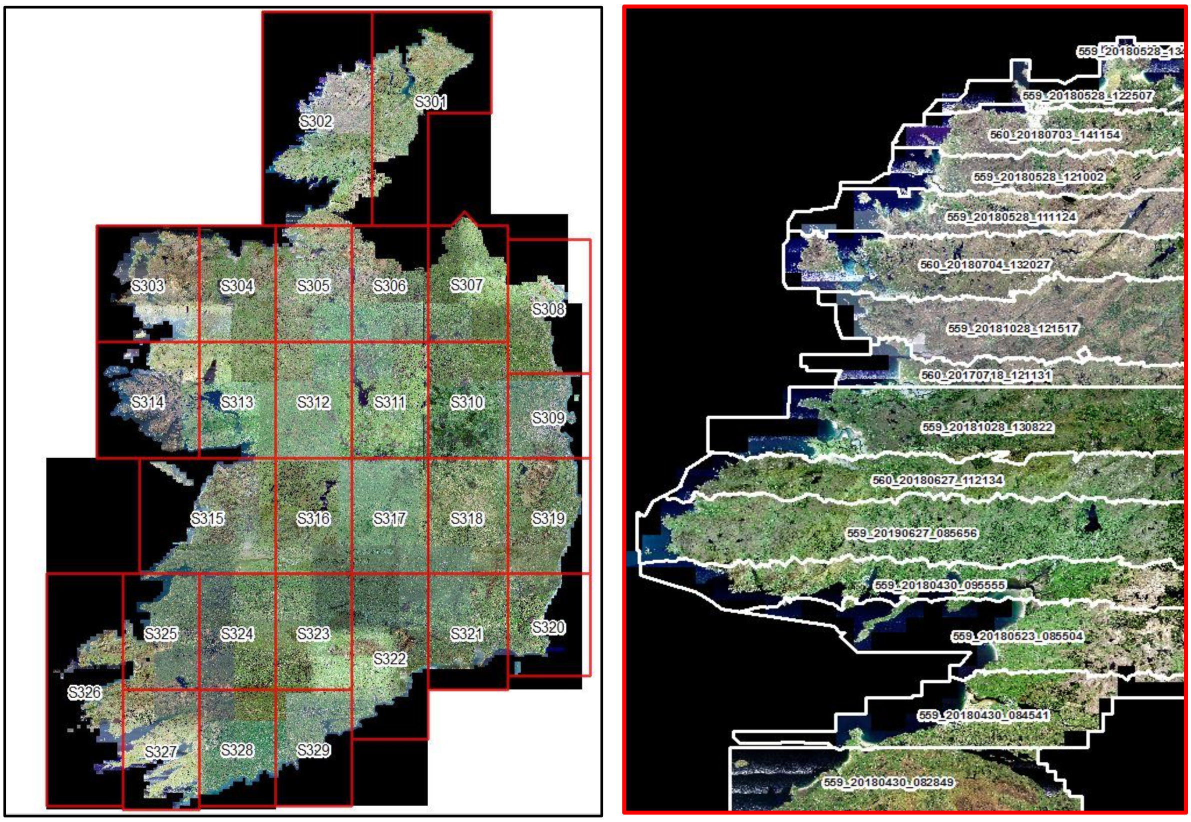


Figure 2.1: Overview of the working blocks in the left-hand image with a close up of block 02 in the North West area on the right showing the image seamlines.

2.3.4 Software and Systems Testing

This project involved the processing, integration and analysing of large complex vector and raster GIS datasets. A process of trialling and testing the available GIS and remote sensing software was undertaken to ensure the capability in processing high-resolution imagery and other core datasets to produce the desired data outputs. A competitive tendering process was undertaken to find the most suitable software packages for the project with a list of key requirements for data processing, integration and classification capabilities stipulated.

The following range of software's were selected to undertake various tasks within the workflows.

GIS Data Processing and Integration

- FME 2018.1 & 2021.2
- ARCGIS Pro 2.6.0
- ArcMap 10.8.1
- QGIS 3.10.5 & 3.20.3

EO Image Processing and Classification

- ERDAS Imagine 2016
- Trimble eCognition 9.3 & 10.0.2

Validation Analysis

- R v4.0.2 & RStudio v1.1.447

2.3.5 Revision of Classification System and Technical Methodology

Following the data collation phase above, it was clear that both the classification system and technical methodology developed by the NLCHM working group required revision to take full advantage of the new, high resolution data streams available to the project. Both reviews were undertaken in parallel with each other as the classification detail achieved and the technical approach used were dependent on each other. Chapters 3 and 4 detail the final classification system and methodological approaches respectively used, with both requiring further revision during the implementation of the project to address emerging issues and limitations based on feedback from validation activities.

2.3.6 Project Management

Several other project management tasks were carried out during the project setup phase. These are outlined below.

Team Resourcing and Training Plan

The necessary skills and experience required to successfully deliver each task was drawn up into a project skills matrix with staff from both agencies matched against them accordingly. This helped to identify any gaps in expertise which informed the development of a training plan for staff in the GIS and Remote Sensing software to be used in the project. Where there was a clear gap in staffing resources, new staff were recruited where possible to fulfil the required role or consultancy support services were acquired through a competitive tendering process.

Project Meetings & Communications

A communications plan was developed both for internally within the project team, with senior management in the OSi and EPA and with external stakeholders. As the team were working across different agencies with different Information and Communication Technology (ICT) systems and protocols, common platforms and communication channels were setup. Weekly team meetings and daily technical sessions were held remotely via Microsoft Teams or Zoom video conferencing software. Regular on-site meetings were also scheduled in the OSi and EPA offices at the start of the project to enable team building and training exercises. Following the onset of the COVID-19 pandemic, all work activity and team engagement moved online with staff working from home.

Stakeholder Engagement

During the development of the proposal for a national land cover mapping programme, a wider network of stakeholders was developed, including experts in topic areas such as agriculture, forestry, peatlands, etc. These experts were regularly contacted for guidance and assistance throughout the project and this support was always generously provided to the project team.

Project Monitoring Tools

A series of Key Performance Indicators (KPIs) and other metrics were agreed to track the project's progress and ensure that satisfactory progress was being maintained. The working blocks and biogeographical areas were used to track the progress of the project across the country. A series of internal and external quality assurance steps were introduced into the production plan to ensure that the data meet the required standard, this is detailed in section 5.

Finalised Project Plan

Following the various testing and systems development activities in the project setup phase listed above, a final project plan was defined considering the design scope and data and software processing requirements.

3. Land Cover Classification System

To date, there has been no recognised national land cover classification system in Ireland. Most studies aiming to map, describe or report on the landscape have used classification systems such as the Pan-European CORINE land cover classification system (Kosztra & Büttner, 2019), the Heritage Council's Guide to Habitats (Fossitt, 2000), or their own bespoke system such as the Teagasc 1995 Land Cover Map (Green, 1995) and OSi's PRIME dataset which has attribution on 'Form and Function' which relate to land cover and land use. This fragmented approach has led to inconsistent mapping and reporting on the landscape and difficulties in cross-comparison of different reports and studies.

A key first-task in developing the National Land Cover Map of Ireland was to design a Land Cover Classification System that could be used in the project to adequately describe the various land surface types in Ireland. Another aim was to design a system that would provide a common standard of land cover descriptions that could be used in other studies to ensure consistent and comparable reporting on the Irish landscape and environment.

3.1 Development of a Prototype Classification System in NLCHM group

Prior to the commencement of the NLC 2018 mapping project, there was initial work done by the NLCHM group to develop a draft Irish classification system and methodology. A desk study (Lydon, 2016) was undertaken by the cross-agency group to research and propose a new national Land Cover Classification System for Ireland. This involved looking at the various approaches and best practices to land cover classification across Europe, assessing stakeholder needs for land cover information in Ireland and proposing a classification system that would balance meeting those user needs with the technical capabilities and data available at that time.

3.1.1 Review of Land Cover Classification Systems in Europe

A review of current national land cover systems in Europe showed that the topic of land monitoring, mapping and classification was very dynamic, with new trends in data sources, technologies and consequently, classification approaches.

Early Land Cover / Land Use Classification Systems

Previously national and pan-European land cover classification systems developed in the 1990's used what were known as 'hybrid' classification system which included a mix of both land cover and land use classes and definitions. This was done primarily to meet user needs which typically require a combination of both land cover and land use information, it also suited manual photo-interpretation analysis which was commonly relied on in earlier systems. The most widely-known hybrid classification system is that of CORINE which has detailed land use related artificial classes such as class 121 'Industrial or commercial units' and class 124 'Airports' but also has land cover classes such as class 412 'Peat Bogs' and class 311 'Broadleaved Forest'.

Pure Land Cover Systems

Distinct 'pure land cover' classification systems then emerged where the classes and their definitions were more solely focused on the bio-physical material on the land surface and its spectral reflectance properties. This approach enables more automated remote sensing classification without the complication of land use activity which is not always discernible from EO imagery. It allows for more accurate and independent class names and definitions and avoids sectoral bias or sensitives around land use reporting.

The classification system adopted by the UK in their Land Cover Map (LCM) 2007 (Morton et al, 2007) and subsequent LCM datasets is an example of a pure land cover approach. In this they took the existing UK's Joint Nature Conservation Committee (JNCC) Biodiversity Action Plan Broad Habitat types (Jackson, 2000) and adjusted the LCM 2007 classes (see Figure 3.1) to suit mapping via remote sensing. For example, the system converted classes such as 'Broadleaved, Mixed and Yew Woodland' into a simpler 'Broadleaved woodland' class which was more achievable to map using remote sensing.

Aggregate class	Aggregate class number[1]	Broad habitat	LCM2007 class	LCM2007 class number[2]
Broadleaf woodland	1	'Broadleaved, Mixed and Yew Woodland'	Broadleaved woodland	1
Coniferous woodland	2	'Coniferous Woodland'	'Coniferous woodland'	2
Arable	3	'Arable and Horticulture'	'Arable and horticulture'	3
Improved grassland	4	'Improved Grassland'	'Improved grassland'	4
Semi-natural grassland	5	Rough Grassland	Rough grassland	5
		'Neutral Grassland'	'Neutral Grassland'	6
		'Calcareous Grassland'	'Calcareous Grassland'	7
		'Acid Grassland'	Acid grassland	8
		'Fen, Marsh and Swamp'	'Fen, Marsh and Swamp'	9
Mountain, heath, bog	6	'Dwarf Shrub Heath'	Heather	10
			Heather grassland	11
		'Bog'	'Bog'	12
		'Montane Habitats'	'Montane Habitats'	13
		'Inland Rock'	'Inland Rock'	14
Saltwater	7	Saltwater	Saltwater	15
Freshwater	8	Freshwater	Freshwater	16
Coastal	9	'Supra-littoral Rock'	'Supra-littoral Rock'	17
		'Supra-littoral Sediment'	'Supra-littoral Sediment'	18
		'Littoral Rock'	'Littoral Rock'	19
		'Littoral Sediment'	Littoral sediment	20
			Saltmarsh	21
Built-up areas and gardens	10	'Built-up Areas and Gardens'	Urban	22
			Suburban	23

Figure 3.1: UK LCM Classification system, showing the link to the UK Broad Habitats system

Pure land cover information on its own does not meet all user needs however, and many applications need to use it in combination with some degree of land use information. The Land Information System of Austria (LISA) system (Banko et al. 2010) developed a twin land cover and land use classification system and datasets, allowing for fully automated classification of land cover type from EO imagery with classification of the land use dataset by integration of raster and in-situ vector data. The Spanish SIOSE land cover programme (Valcárcel et al., 2008) also has a dual land cover and land use system, where each geometry can only have a single land cover code comprising 100% of the geometry whereas there can be multiple land use class types for a single geometry with different area percentage values given.

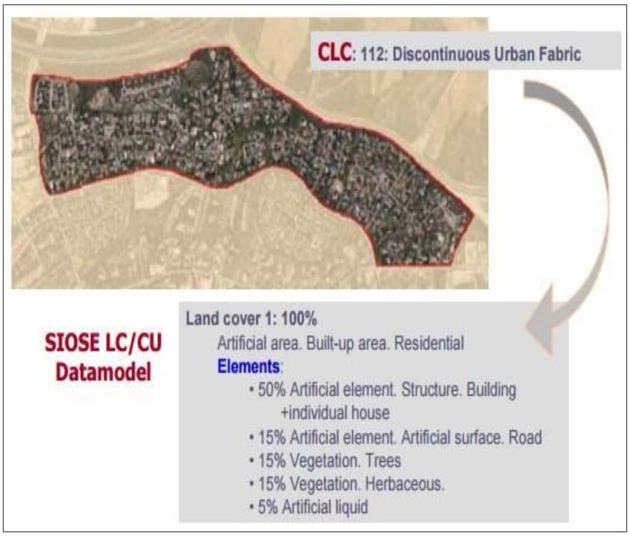


Figure 3.2: Example showing the Spanish SIOSE classification approach where a primary land cover class is given along with percentage cover of several landscape 'Elements'.

Landscape Description Model

These were the first steps towards a new disaggregated and modular approach to national land monitoring and classification called the 'landscape description model'. Instead of trying to satisfy all user needs within one hierarchal system, these systems broke the landscape into its different constituent parts in separate but related data tables. The EIONET (Environment Information and

Observation Network) Action Group on land monitoring in Europe (EAGLE) developed a Pan-European landscape description system called the 'EAGLE matrix' (Arnold et al., 2020). This was a flexible but comprehensive land classification system which has a separate land cover and land use tables with also a third group of tables covering additional attribution that can be populated with available land information where available (see Figure 3.3 below). The use of the Landscape characteristics module allows the capacity to store and catalogue valuable information on the landscape such as agricultural land management practises, forestry planting dates, conservation status, population records etc. This complementary information, which could not be accommodated in the more rigid land cover and land use hierarchies, can be referenced, and integrated with the land cover and land use data as needed to answer specific user and policy data needs.

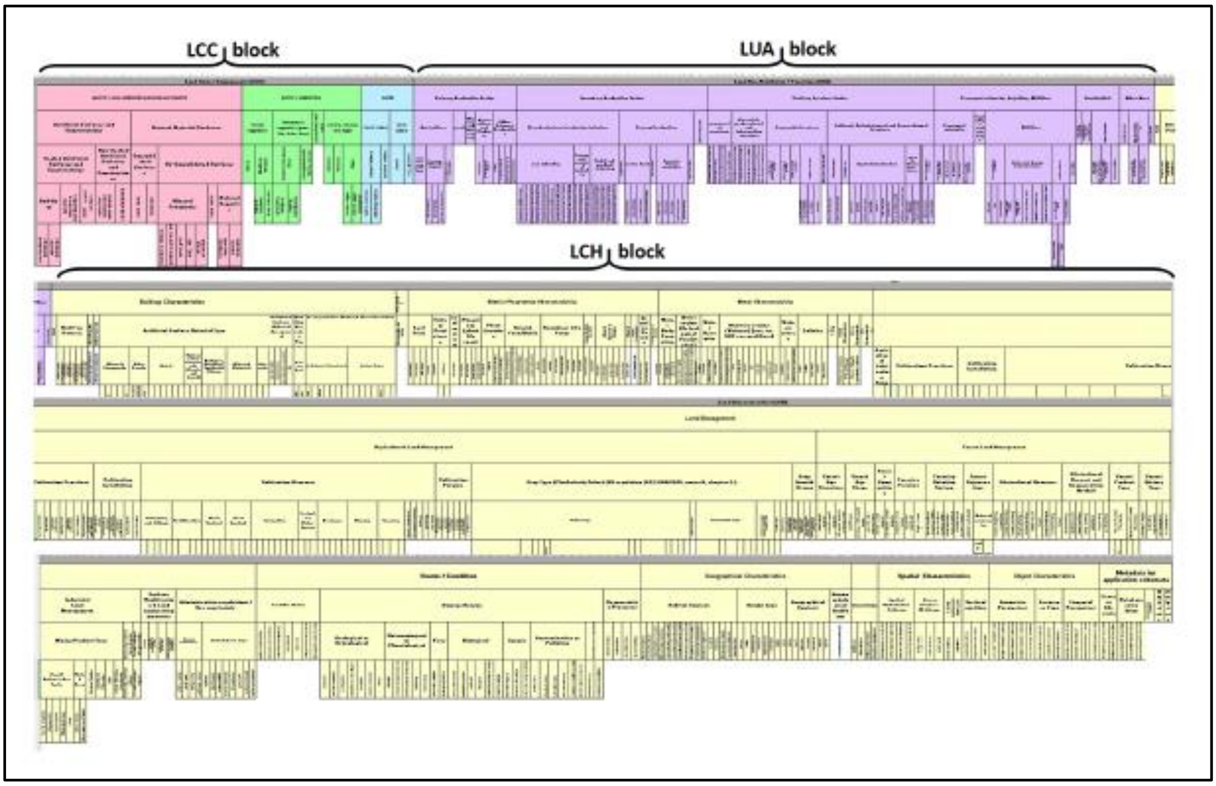


Figure 3.3: Overview of the EAGLE matrix with separate modules for Land Cover Components (LCC), Land Use Attributes (LUA) and Land Characteristics (LCH) – click here for official EAGLE matrix

CORINE+

The landscape description model approach has recently been implemented in the new CORINE+ system, also known as CLC+. The aim of CLC+ is to provide a modern spatial infrastructure for capturing land cover, land use and land characteristics at a better spatial and temporal resolution. CLC+ is part of the Copernicus Land Monitoring Service (CLMS) and it will provide the basis for a range of pan European products into the future. This will include a second-generation version of the traditional CORINE land cover dataset, known as CLC+ Legacy, which will provide continuity with the CORINE data time series.

The CLC+ spatial data infrastructure will include a grid-based data model which ingests a wide array of land data products from member states and uses different combinations of the land cover, land use and land characteristics described above to derive tailor made data products called 'Instances'. Examples of CLC instances include European-wide LULUCF, CAP and Urban Atlas layers.

3.1.2 Developing a Land Cover Classification System for Ireland

Looking at the best practise in Europe when developing a new Land Cover Classification System for Ireland, and bearing in mind there was a user need for land use, habitat and other land data in addition to land cover information, the NLCHM group decided to aim towards developing an Irish landscape description system, with a land cover classification schema being the first step in that process.

Adapted Fossitt Level 2

As with the UK approach, it was logical to try and align the classification system with the established habitat classification system used in Ireland - the Heritage Council's Guide to Habitats in Ireland (Fossitt, 2000). Pilot studies using the data available at the time showed it was possible to map most of the classes at the existing Fossitt Level 2 which equates roughly to the standard level of land cover detail in other countries. The study then sought to adapt the existing Fossitt Level 2 to one that was fully achievable using automated remote sensing methods and a draft new national land cover system was drawn up.

Stakeholder Engagement

Widespread stakeholder engagement was undertaken as part of the development of the system with experts from across the different environmental topic areas such as Agriculture, Forestry, Peatlands and Grasslands being consulted through technical workshops and an online stakeholder survey to assess their data needs in this area and gather feedback on the proposed approach. The survey also tried to scope the needs or requirements of responders for additional types and levels of data to provide a full picture of the needs of Irish land data users.

Over one hundred responses were received from environmental experts and data users representing sixty individual public agencies, companies, academic researchers, and NGOs. Most survey responders were satisfied with the proposed classification approach with 70% of users saying that they agreed with this approach. Although over 80% of responders said that they would use the data if produced, in some land cover topic areas there were requests for more detailed thematic description and classification.

Based on the satisfaction of the user groups to the proposed approach, the review recommended that an adapted version of Fossitt Level 2 should be used as the basis of a national land cover classification system, but that this should be reviewed when the project commences and revised if necessary based on the data and technologies available at the time.

3.2 A New Land Cover Classification System for Ireland

As mentioned in section 2, the land cover project had access to the OSi's recently captured 2018 15cm (Dublin) and 25cm (rest of country) ortho-photography and 1m Digital Surface Models (DSM), along with 10m dual season Sentinel satellite imagery and the national cadastral vector dataset – PRIME2. This was a significant increase in data detail and quality from when the pilot work, which was undertaken based solely on Sentinel 2 imagery. Access to better data represented an unprecedented opportunity to map the Irish landscape at a very high level of spatial and thematic resolution. A review of the classification system was then undertaken to investigate if more detailed thematic and spatial detail could be achieved using the new data sources.

3.2.1 Revision of Classification System for the Land Cover Project

The revision work initially investigated whether the new data allowed mapping of all Level 3 habitats in the Fossitt system. Whilst some Level 3 classes could be mapped, including buildings, hedgerows and all five classes in the peatland group, more classes could still only be mapped at the adjusted Fossitt Level 2 such as Broadleaved Woodlands, Exposed Rock and Cultivated Land. Translating the more detailed Level 3 Fossitt definitions and rules into a remote sensing-based classification was also problematic, as the class definitions were based on habitat systems and the presence of indicator species and environmental context, which are not always discernible from automated analysis of high-resolution aerial photography. This meant it was very difficult to definitively say whether the classification was meeting the definitions used in the Fossitt system or not. As this project was not using ground-based sample data, classified to the Fossitt system, to train the classification models, we could not rely on these to train and validate the classification outputs.

Therefore, it was decided that it would be more logical to develop a new classification system, still aligned where possible to the Fossitt system, but designed to better suit remote sensing-based analysis, using criteria and definitions based primarily on interpretation of satellite and aerial-based earth observation imagery and readily available in-situ vector data.

3.2.2 Land Cover Classification System for Ireland

The new Land Cover Classification System of Ireland is presented in Table 3.1 below. It is an applied remote-sensing based classification, meaning that as described above, the class definitions are based on the technical approach to mapping them using present day remote sensing and data integration techniques.

Class Rules

Five rules were used as a framework to determine which classes or level of detail to be used in the system:

1. All classes should be primarily defined by their bio-physical properties and the spectral characteristics of these properties (i.e. how they appear) in Earth Observation imagery.
2. Where classes or definitions are used that are not based on spectral and/or height values in EO imagery, authoritative, national-scale and readily available in-situ vector data should exist to enable accurate classification.
3. All classes should be mutually exclusive, this means that the same definition criteria should not be used in more than one class and any single spatial object in the dataset / landscape should belong to one class only.
4. Class definitions and terminology should be independent of any sectoral bias and a balance in detail should be sought across all thematic sectors and land cover types.
5. All classes should be practical to validate, i.e. can be done using data and systems that can be distributed to validators and not rely on large-scale ground surveying.

Following several iterations of testing and development, supported by BEC consultants who were providing ecological support to the project, a list of spectral indicators and definitions were developed. Parameters such as spectral tone, smoothness and texture, height and slope values, the presence of shade and seasonal variation were used to provide a coherent and mutually exclusive set of classes to adequately describe all land cover types in Ireland.

As the full range of land cover types in Ireland had not being mapped in this detail before, it was also necessary to allow for some adjustment and revision of the system throughout the production process. This was based on internal validation results and feedback from the steering and external validation partners.

Structure of the Classification System

The Land Cover Classification System is a two-level hierarchal system giving general land cover groupings at Level 1 with more detailed classes at Level 2 (see Table 3.1). All Level 1 classes should be readily mapped from freely and frequently available EO imagery. Level 2 classes are more challenging to map and some rely on high resolution aerial imagery and / or in-situ vector data. This system is designed based on current capabilities and knowledge of the Irish landscape. This is an evolving topic area meaning the system should be reviewed and updated, if necessary, in line with improving data and technical capabilities.

Code	Level 1	Code	Level 2
100	ARTIFICIAL SURFACES	110	Buildings
		120	Ways
		130	Other Artificial Surfaces
200	EXPOSED SURFACES	210	Exposed Rock and Sediments
		220	Coastal Sediments
		230	Mudflats
		240	Bare Soil and Disturbed Ground
		250	Burnt Areas
300	CULTIVATED LAND	310	Cultivated Land
400	FOREST, WOODLAND AND SCRUB	410	Coniferous Forest
		420	Mixed Forest
		430	Transitional Forest
		440	Broadleaved Forest and Woodland
		450	Scrub
		460	Hedgerows
		470	Treelines
500	GRASSLAND, SALTMARSH and SWAMP	510	Improved Grassland
		520	Amenity Grassland
		530	Dry Grassland
		540	Wet Grassland
		550	Saltmarsh
		560	Sand Dunes
		570	Swamp
600	PEATLAND	610	Raised Bog
		620	Blanket Bog
		630	Cutover Bog
		640	Bare Peat
		650	Fens
700	HEATH and BRACKEN	710	Bracken
		720	Dry Heath
		730	Wet Heath
800	WATERBODIES	810	Rivers and Streams
		820	Lakes and Ponds
		830	Artificial Waterbodies
		840	Transitional Waterbodies
		850	Marine Water

Table 3.1: The new Land Cover Classification System of Ireland

Although a 'pure land cover' system was the most preferable systems from a remote sensing perspective, in practise, adhering strictly to this rule was not always possible. This was due to a strong stakeholder desire for specific land use related classes or because land use is an inherent element of the land cover types in Ireland and it would not be logical to separate them. The three instances of this are Class 120 Ways, Class 310 Cultivated Land and Class 520 Amenity Grasslands. All three classes rely on land use information obtained by in-situ data for complete classification.

Class 630 'Cutover Bog' infers a historical land use and was included as it is a critical national land cover type, identified as such through previous stakeholder engagement. The identification of cutover areas can be inferred from EO imagery but can only be mapped reliably in a semi-automated fashion, using a degree of manual interpretation.

Throughout the development of the classification system there was ongoing engagement with the wider stakeholder network and close consultation with the projects ecological consultants BEC, who provided expertise on this topic to the project. A classification guide is being produced to accompany the dataset in conjunction with BEC, this guide will detail the definitions and criteria used to map each individual class. Summary definitions for all classes are also given in Appendix 1.

4 Data Processing and Classification Methodology

A multi-stage production methodology was developed to process, analyse, and classify the range of raster and vector-based data sources collated for the project. This chapter details these processes and the different stages of the technical methodology. It also describes the internal validation steps used to ensure that the data products were meeting the required standards and the post-processing steps used to create and refine the final national scale data product.

4.1 Data Pre-Processing

Before the classification process could begin, the various input datasets had to be processed to remove errors, transform them into a standardised format and make them 'analysis-ready'. The various pre-processing steps are outlined below, separated into vector and raster data types.

4.1.1 Pre-Processing of Core Vector data

The core vector datasets used in the overall classification process are:

- PRIME 2
- LPIS
- NPWS coastal habitat data
- Forest data – including Forest Service, Coillte & NPWS woodland data
- Seamlines dataset

The PRIME2 dataset was chosen to provide the baseline spatial framework for the classification process. Its spatial boundaries were then used as the spatial 'containers' for the image segmentation and classification processes. All other core vector datasets above, except for the Seamlines dataset, were then processed, aligned with, and integrated into the PRIME2 dataset, using FME data transformation tools, and involving the following steps:

Geometric Cleaning

All data was checked for geometric and topology errors, duplicate polygons and matching neighbouring polygons that were dissolved together to reduce the data size.

Attribute Cleaning

The land cover, land use or habitat attribute data contained within the source data was analysed and then translated or re-categorised into a temporary Intermediate Land Cover Classification (ILC). This was a list of general land cover groups which were used to inform the spectral classification process later in the workflow. Where there was a one-to-one relationship between the source data attribution and a class in the new Land Cover Classification System, then the in-situ data field was assigned to the relevant land cover class. An example of this would be all arable crop records within the LPIS database which were directly classified as Class 310 'Cultivated Land'. Other source data which had a one-to-many, or many-to-many relationship with the classification system were given an ILC label. An example for this would be the LPIS entries such as 'Rough Grazing', 'Permanent Pasture' and 'Traditional Hay Meadow' which were all given the ILC label 'Grasslands'. Such parcels were subsequently classified into their relevant Level 2 grassland category based on the spectral values in the imagery.

Data Integration

The cleaned source data was then spatially integrated into the master PRIME2 dataset. The source data was firstly snapped to align its spatial boundaries to that of PRIME2, then using spatial overlap tools the geometries and attributes of the source data were transferred to the PRIME2 dataset using overlap and area thresholds. For example, if a LPIS parcel had a >90% overlap with a PRIME2 field boundary then they were considered to be the same and only the attribute information was transferred to the PRIME2 polygon. For overlap <90% a 0.1ha threshold was set so that any geometry greater than this and their attributes were transferred over to the PRIME2 dataset. A series of geometric cleaning operations were run, after the integration process, to ensure that any gaps, overlaps, etc. were not imported into the final dataset.

The resulting dataset from the above processes was an enhanced-PRIME2 dataset, its original attribution, and geometries untouched but with additional geometries and land use and environmental attribution. This was then imported into the classification software to inform the image segmentation and classification processes.

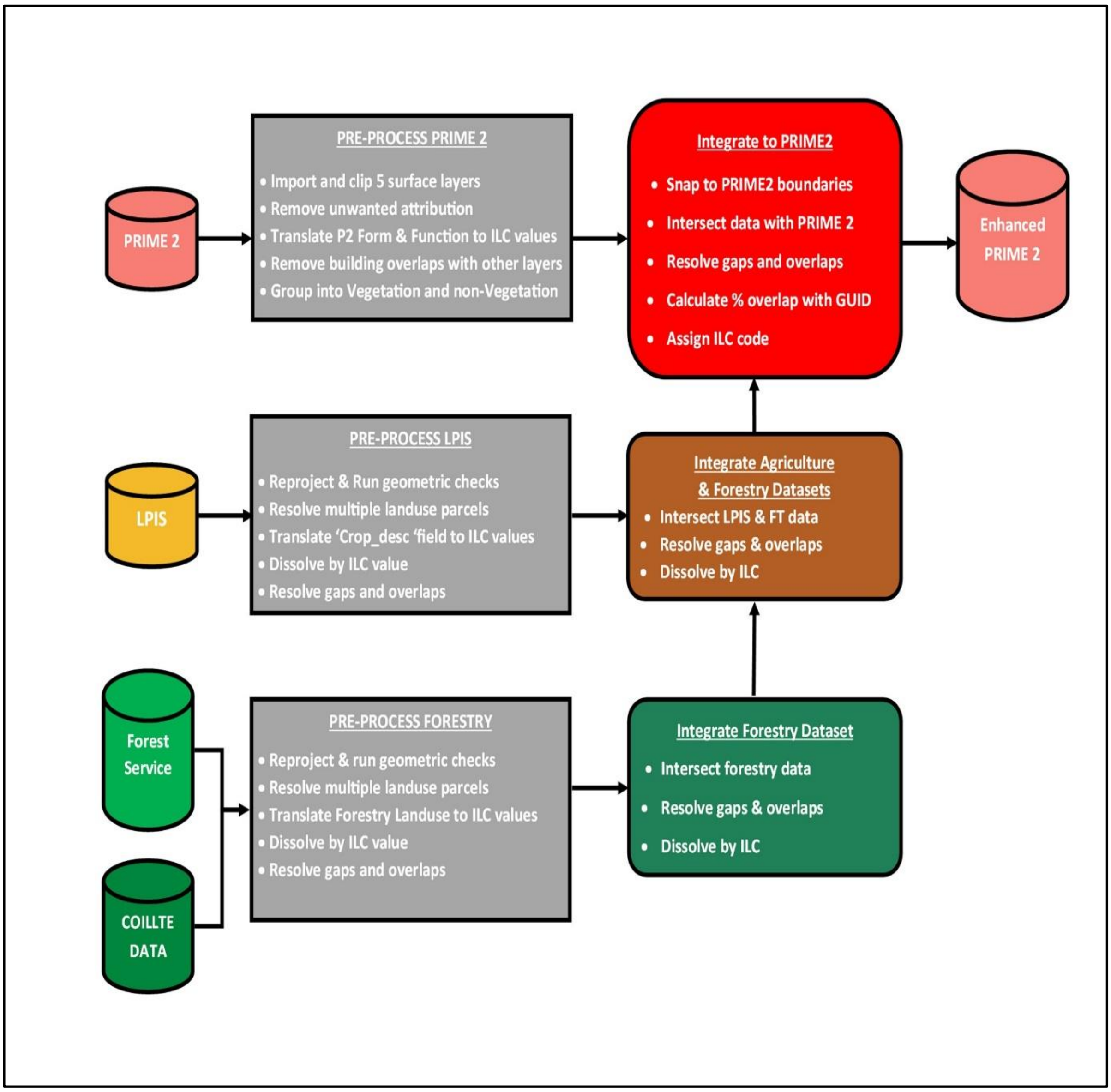


Figure 4.1: Pre-processing core vector data integration workflow

Seamlines Dataset

The seamlines dataset was used to delineate individual classification scenes based on the spatial extent of the source ortho image strips. Adjacent seamlines which were captured on the same flight and so had the same capture date, were merged together and given a unique 'Tile_ID' to allow the eCognition ruleset to identify and split the imagery based on the tile outlines.

Secondary Vector Data

The secondary vector data which was to be used as visual reference data was kept in its native format and used in most cases to provide environmental context and to aid the selection of sample objects for the machine learning classification. It was not spatially integrated with PRIME2, but standard geometric cleaning was undertaken to enable smoother referencing and rendering.

4.1.2 Pre-Processing of Raster Data

Like the vector data, all raster datasets required several pre-processing steps to prepare them for automated remote sensing analysis and classification.

Series 2 Ortho-Photography:

The Ortho-photography was initially supplied in 25cm resolution, 15cm in Dublin, as 4-band (R, G, B, NIR) 2km² tiles. After initial testing with the 25cm .tif data it was decided to resample the imagery to 1m pixel resolution and into the .img format as this reduced the data volume by 75%, significantly improving computation times with a minimal reduction in image detail. The 2km ortho tiles were merged and clipped to the full working block areas and imported for analysis into eCognition.

Series 2 Digital Terrain Model (DTM) & Digital Surface Model (DSM):

As with the Orthos, the OSi Series 2 DTM & DSM were each supplied in 2km tiles which were then merged to create a single .img raster file for each working block. The source resolution of the DTM was 2m and the DSM was 1m. The DTM was processed to generate a slope raster and contour vectors at 10m intervals in QGIS. As this was a photogrammetry based DTM, there were some data gaps in areas of high slope relief and in large forestry blocks. Image pixel filtering techniques were used to infill the gap pixels which had negative or no data values.

Normalised Difference Surface Model

A 1m Normalised Surface Difference Model (NSDM) was created by subtracting the values of the DTM pixels from those of the DSM to produce a raster with real above-ground height pixel values.

Image Filtering

All high-resolution raster images were put through several convolutional and median filtering processes to remove negative or erroneous data values, spectral noise caused by unusually high or low reflectance values. This allowed a smoother image to be produced for the segmentation and classification processes that followed.

Sentinel 2 Imagery

Sentinel Imagery from Spring and Summer 2018 was utilised in addition to the ortho photography to identify temporal variations in many classes. The Sentinel imagery was downloaded from the ESA Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus/#/home) for all the Sentinel Grid areas that intersected the set of working blocks. Level 2A imagery was downloaded where available for all Sentinel Grid areas. Where 2A level imagery was not available 2C level data was downloaded and ran through Sen2Cor command in QGIS to create level 2A imagery locally.

A Spring and Summer layer stack of images was created for each Sentinel Grid area, clouded and shaded pixels were masked out and filled in with pixels from the nearest available date. The composite images were then reprojected to the ITM projection and clipped out to the working block areas.

4.2 Land Cover Data Classification

Following all the above pre-processing steps the core source data was ready to be loaded into the image classification software, Trimble eCognition, for object-based image analysis.

4.2.1 Object-Based Image Analysis (OBIA)

OBIA is distinct from traditional pixel-based classification where individual image pixels are the spatial unit of classification. In OBIA, a two-step classification process is used involving image segmentation followed by object-based classification.

Image Segmentation

During image segmentation, neighbouring image pixels with similar spectral values are grouped together using image clustering techniques to create 'image-objects'. These objects are similar to vector polygons and delineate real-ground features visible within the image. The objects are then considered as a single entity and used as the primary spatial unit of classification instead of individual pixels. This allows for the use of aggregated spectral band statistics in the classification process such as mean, median, min / max, and standard deviation along with shape and geometry-based values e.g. area, length, width, and border index. There are several different types of segmentation available including vector-based segmentation which replicates the geometries of input vector data, multi-resolution segmentation which is based on the spectral values of the selected image bands and chessboard segmentation which creates grid-like objects set to a desired pixel width.

Object Classification

Following segmentation, remote sensing classification approaches including supervised and machine learning algorithms, can be used to assign a class based on a set of training samples which can be either imported or collected directly within the software from the source imagery. Hard rules can also be written into the workflow to directly assign a class based on band values, in-situ vector information (e.g. agri-land use datasets), geometry values (e.g. object area & width), neighbourhood & proximity functions (e.g. shared borders and distance). Often a combination of both hard rules and supervised classification are used in a logical sequence to obtain the best results. This requires more effort in the development of the classification methodology, but it allows for a more knowledge-based decision making (Blaschke, 2008) and less reliance on black-box classification algorithms.

4.2.2 Classification Methodology

An OBIA classification algorithm was then developed to produce a vector classification product for the whole country based on the data available and the newly designed Land Cover Classification System described in chapter 3. Developing the algorithm involved several testing and refinement cycles to identify the most efficient workflows to process and classify the large quantities of data used and to produce the best quality outputs possible.

The main stages of the algorithm are summarised below:

i. In-Situ Based Classification

A number of classes were classified directly from the attribute information and geometries supplied by the PRIME2 vector layer, as it was deemed that it would not be possible to produce the same level of mapping detail for these classes directly from the imagery. These were the Buildings, Ways and Waterbodies classes, which were then classified at the PRIME2 level based on the FORM and FUNCTION attribution within the PRIME2 data. Separate to this, the NPWS coastal habitats data was also used to directly classify PRIME2 GUIDs which contained Sand Dunes, Saltmarsh, Mudflats and

Transitional Waterbodies. The Cultivated Land class was classified directly from the LPIS dataset, while forest land recorded within the in-situ forestry land use data was classified based on species information in the in-situ data combined with the NSDM height values.

ii. Segmentation and Classification of Unmapped Forest Lands

The forest land areas not included in the in-situ data was referred to as 'un-mapped' forest lands and these areas were classified prior to the other land cover classes due to the unique complexities in mapping this group such as the influence of height, shadow, in-situ data and seasonal spectral variance.

Unmapped Forest Lands were dominated by broadleaved forest, linear woodland, and scrub. A bespoke rule-based algorithm module was developed to map these based on height and width thresholds, in addition to spectral values and spatial relationships to the PRIME2 boundaries. Special modules were developed to mask out and remove shadow which was a prominent feature in the Orthophotography and also to remove false height values in the NSDM which occurred along the perimeter of forests and hedgerows due to the orthogonal slant of the camera lens.

iii. Machine Learning Classification

All other classes were classified using the random forest machine learning classification algorithm which has proven to produce good results in mapping land cover types in Ireland (e.g. Cawkwell et al. 2016). This involved the following steps:

- *Sample Selection*

  A sample selection sequence was developed whereby the interpreter was presented with a tranche of randomly selected image objects which they manually classified to the Land Cover Classification System. These samples were then split 80/20 per class meaning 80% of the samples for each class were used for training and 20% for validation. As no suitable reference dataset was available to stratify the samples based on class area, an unstratified sample collection strategy was used with a number of rules enforced to ensure a representative number of samples were collected for all land cover types present within the scene. Due to the image data heterogeneity it was also important to ensure that samples were selected across the entire classification scene.

  The number of samples selected was based on the total area of the classification scene, with one sample collected per 1km2. Areas with known high land cover variability (e.g. uplands, coastal and designated areas) were weighted to ensure a higher sampling rate of 2 samples per 1 km2. To ensure that the samples were distributed across the classification scene, a minimum distance of 250m was enforced between random samples. A minimum of 20 samples per class were collected and no class could represent more than 20% of the sample set.

- *Round One Classification*

  The training samples were then used to train the random forest classification algorithm and the model was then run on all unclassified image objects, resulting in a fully classified scene.

- *Internal Accuracy Assessment*

  An internal accuracy assessment was run using the pool of validation samples with an error matrix produced to give the accuracy statistics from the classification process. This assessed how well the classifier performed in reproducing the classification supplied by the sample set, and as such was not an assessment of the ground accuracy of the map product.

- *Round 2 Sample Selection and Classification*
  If the classifier did not perform satisfactorily in any class, i.e. a classification threshold of 75% was not reached, additional samples were collected in these respective classes to try and improve the classification results. Typically, there would be confusion whereby an area of the classification scene consisted of imagery from a different capture date and consequently with a different spectral profile.

  The new samples were entered into the sample set and a new random 80/20 split was performed on each class with a second classification being trained with the new 80% samples. This process was run iteratively until a satisfactory classification accuracy result was achieved across all classes. In most cases two or three sample selection and classification iterations was sufficient to achieve the desired classification accuracies.

- *Sample Transfer*
  The final set of samples were retained and reused to classify other classification scenes within the same working block which had a similar capture date. This greatly increased the efficiency of the classification process as the time-intensive sample selection process was typically only needed on 2-3 seamlines per block.

iv. Data Cleaning and Export

Following classification, a number of processes and rules were run to merge the classified objects together, enforce the various minimum mapping units and manually classify some classes which relied on a combination of in-situ information and the label assigned by the classifier e.g. all improved grassland falling within a PRIME2 artificial area was reclassified to Amenity Grassland. Following this the data was then exported as a shapefile with a single Level 2 class attribution field.

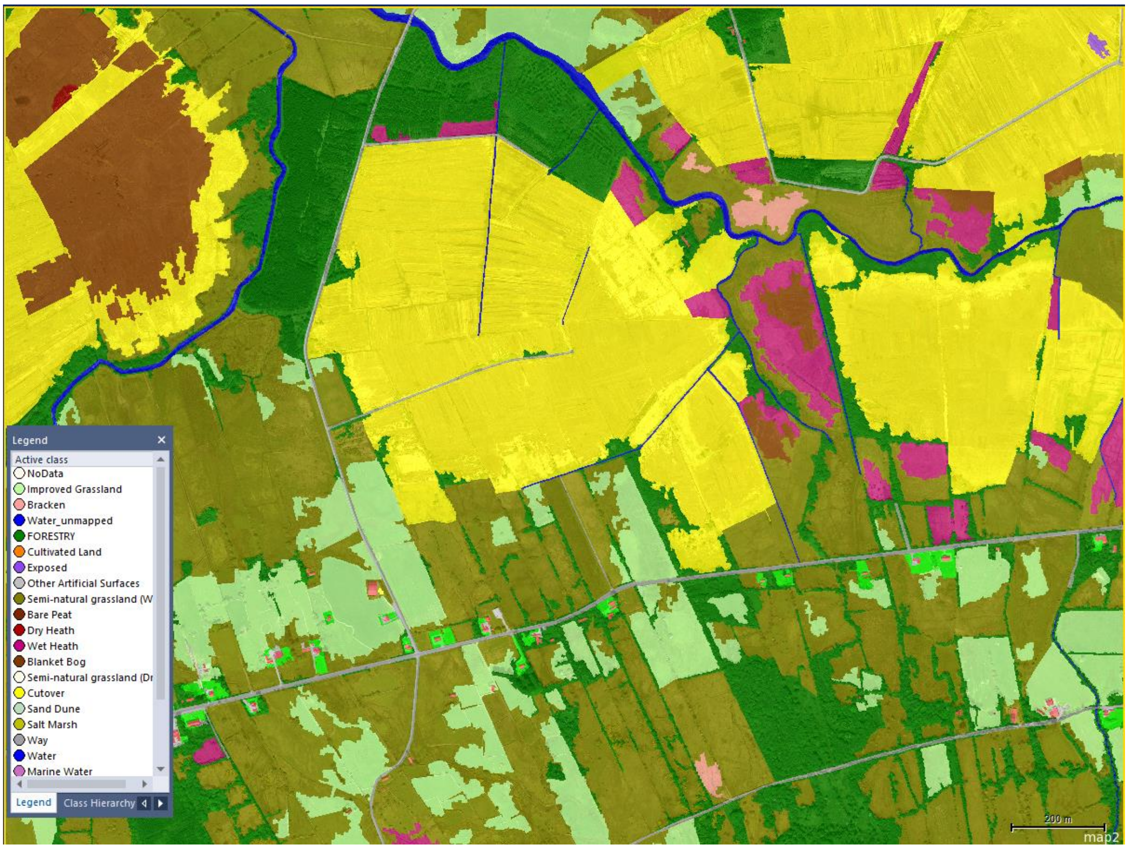


Figure 4.2: Example of classifier output from using object-oriented methodology.

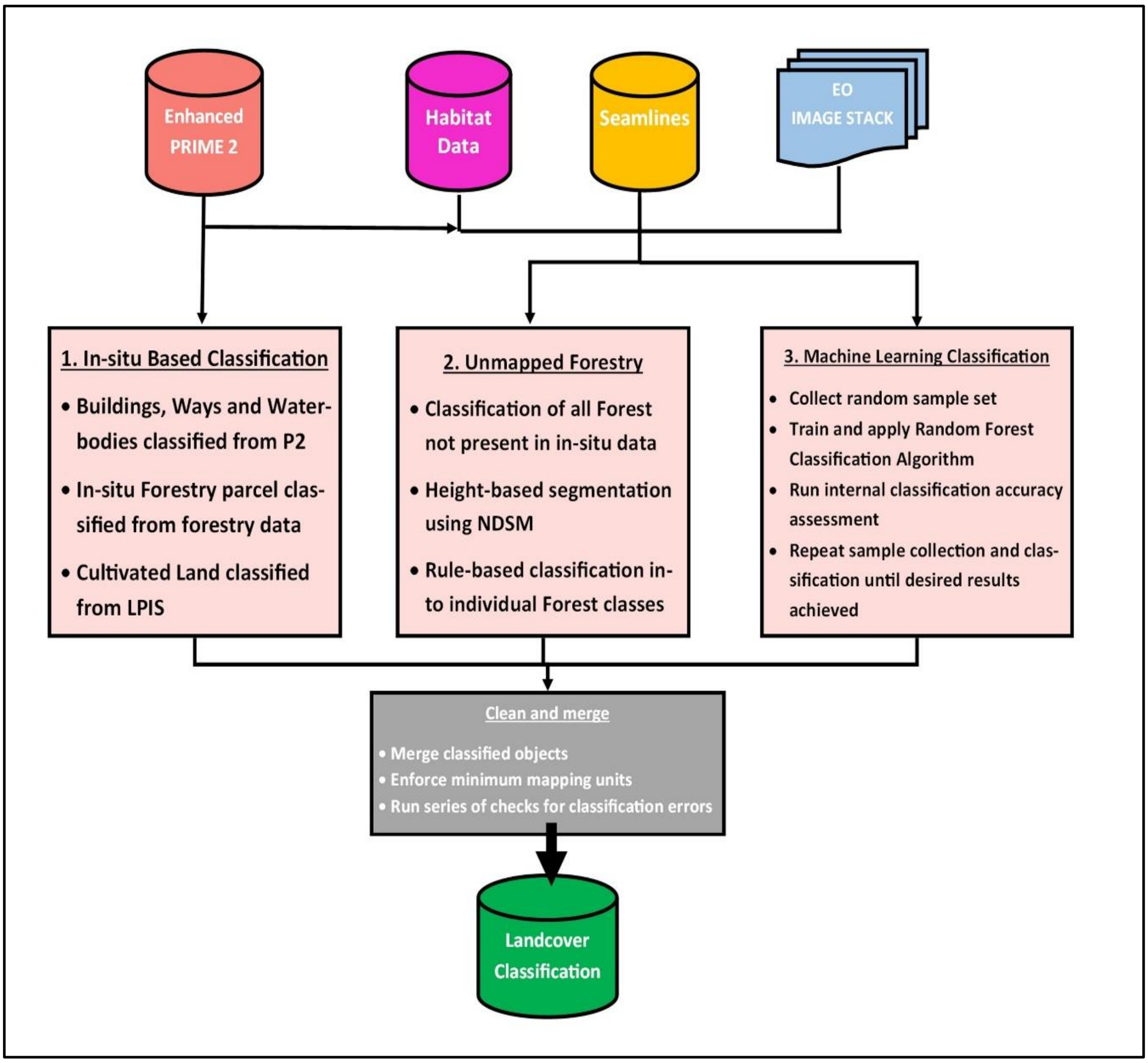


Figure 4.3: Diagram giving an overview of the classification workflow

4.3 Internal Validation

After the classification result was exported, an additional manual internal validation process was undertaken to ensure the data was classified to the required standard before proceeding to post-production. QGIS software was used to generate a minimum of 30 random samples per class which were then manually inspected by the validation team for thematic and geometric accuracy. A quality report was compiled with the results for Overall Accuracy (OA), Omission, Commission, Producer's Accuracy & User's Accuracy for each Level 2 class. A minimum Kappa Index of Agreement, or KIA score, of 75% was required for the scene to pass this stage, any scene falling below that was either sent back for re-classification in the case of systematic errors, or, in the case of isolated errors, manual editing was used to correct the error where appropriate. Following this, the scene was re-validated until the minimum standards were achieved.

4.4 Post Processing

The classified data was then put through a range of post-processing steps to take it from a raw output into the final data product. An overview of the main processes undertaken are shown below:

- *Generalisation:* The data was firstly smoothed and generalised to both remove the pixelated aspect of the classification and reduce the number of vertices.
- *Snapping:* This occurs when a classification parcel is within 5 metres of a PRIME 2 boundary. The classification output then snaps to a PRIME 2 boundary helping improve spatial accuracy and visual clarity.
- *Replacement:* The pixelated buildings / rivers / ways from the classification output are removed and are replaced with the original features from PRIME 2.
- *Integration:* The classified output is integrated with PRIME2 resulting in the PRIME2 geometries being re-introduced to the classified output. Although Prime 2 is not in the final iteration of this product, its presence here ensures a consistent structure is kept as the data is processed, keeping linear boundaries. As the example in Figure 4.3 demonstrates, if a small spike of data needs to be merged into its neighbour, with the Prime 2 GUID it can be ensured it stays within its own boundary and does not appear as a spike in another class.

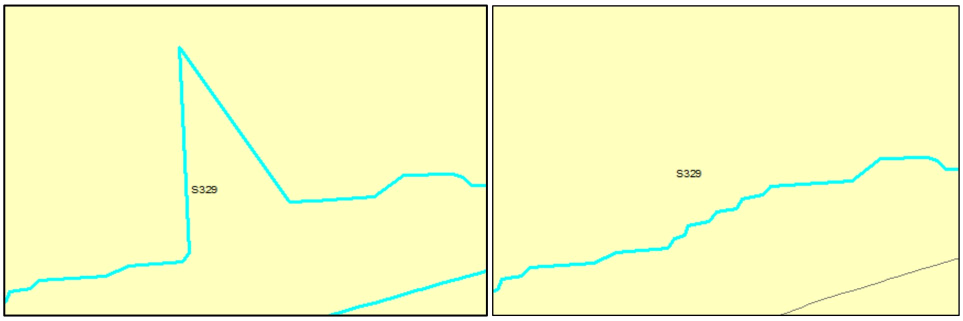


Figure 4.4: Example showing feature snapping without (left) and with (right) PRIME2

- *Minimum Mapping Unit (MMU):* At the classification stage there are a set of class-based MMU's enforced, ranging from 100 - 1,000 sqm. However, in the subsequent post-processing stage this was reduced to 100 – 500 sqm. As the classification and PRIME 2 are integrated during post-processing, and as PRIME 2 has no specific MMU, many smaller geometries are re-introduced to the dataset thus splitting the original classification. This means a second round of MMU enforcement was needed to remove these smaller geometries but the MMU also needed addressing. For example, a 1,998 sqm feature could be split down the middle, resulting in both features falling below the MMU and being dissolved into their relevant neighbour feature, losing valuable information. It was therefore decided to reduce the MMU to 500 sqm to largely avoid this issue.
- *Merging:* The data derived from the classification arrives for post processing in seamline strips. These are first merged into their block components, then into BGRs and finally into the national dataset. Here, the Prime 2 GUID is used to ensure buildings /rivers / streams / ways do not merge into one another. Manual edge matching is necessary to ensure that there are no errors or overlaps along seamlines or block boundaries.
- *Attribution:* The final attribution was added including a new LAND_COVER_ GUID, the flightline date and other attributes as listed in Appendix 2 of this report.

4.5 Final Data Product

4.5.1 Data Format

The final data product delivered as part of the land cover mapping programme is a seamless, national-scale GIS vector dataset, available in file geodatabase (.fgdb) format as both Irish Transverse Mercator (ITM) / ESPG 2157 and Irish Grid (IG) / ESPG 29902 projections.

4.5.2 Minimum Mapping Unit

To maintain a consistent level of data detail and quality in terms of remote sensing classification and land cover mapping, it was necessary to set a minimum mapping unit (MMU) for the data products. The MMU is the minimum size (area or length) of any spatial object in the dataset. Any feature in the landscape which is below the MMU is not mapped and is not considered as an error when omitted from the dataset.

In object-oriented classification, where the classification units are vector objects, a minimum area is chosen based on the resolution of the source imagery pixel and the level of detail pursued. The MMU must always be larger than the pixel size of the input imagery, best practise suggests allowing at least a 5x5 pixel box around an image pixel. For example, for a pixel size of 1m then a 5x5m object window would result in an MMU of 25 sqm.

For this project, the core imagery pixel size was 1m. This would allow for a MMU of 25 sqm. However, mapping at this scale on a national basis would lead to unmanageable quantities of data that are difficult to manage, process, classify and validate. Traditionally a single MMU value is selected for all classes. However, as some classes such as buildings, ways and waterbodies are being derived directly from in-situ vector data, the project team decided to use a series of class based MMU's in order to retain the finer detail provided for some classes, whilst also balancing the practicalities of mapping other land cover classes such as grasslands and peatlands on a national scale. Appendix 3 shows the different MMU sizes for each Level 2 class.

4.6 Data Access

For details on how to gain access to the National Land Cover Map (NLC 2018) please contact corporatesales@tailte.ie

5. National Land Cover Area Statistics

The overall objective of the National Land Cover Map was to provide more detailed and accurate data on the land cover types found in the Irish landscape. This has been achieved through the spatial delineation and classification of over 10.1 million features across the country. This chapter provides the area statistics for all Level 1 and Level 2 land cover classes, measured using GIS analysis of the new NLC 2018 data. It compares these figures to CORINE 2018 and other established land cover and land use datasets to show how the understanding of our landscape will be reconsidered with this new environmental baseline dataset. Examples of the final dataset are also given for key topic areas.

5.1 Level 1 Results

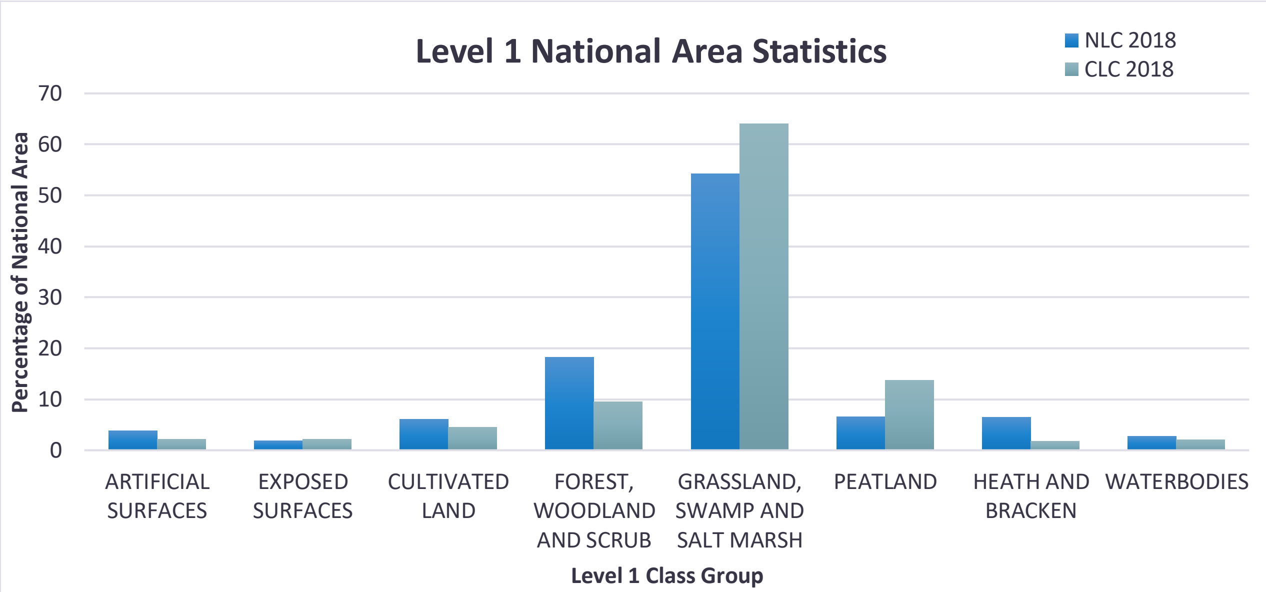


Chart 5.1: Chart comparing the percentage of national area by Level 1 group for the new National Land Cover Map (NLC 2018) and CORINE 2018 (CLC 2018).

	Code	Level 1	NLC 2018 Area (ha)	Corine 2018 Area (ha)	NLC 2018 %	CLC 2018 %	% Diff.
	100	ARTIFICIAL SURFACES	268,016.08	155,961.25	3.79	2.21	+1.59
	200	EXPOSED SURFACES	133,270.48	151,080.74	1.89	2.14	-0.25
	300	CULTIVATED LAND	427,033.09	320,624.53	6.05	4.54	+1.51
	400	FOREST, WOODLAND AND SCRUB	1,290,756.87	672,084.98	18.27	9.51	+8.76
	500	GRASSLAND, SWAMP AND SALTMARSH	3,828,160.69	4,521,366.43	54.20	64.00	-9.81
	600	PEATLAND	462,291.55	968,808.26	6.55	13.71	-7.17
	700	HEATH AND BRACKEN	456,916.50	126,087.98	6.47	1.78	+4.68
	800	WATERBODIES	196,799.81	148,198.01	2.79	2.10	+0.69

Table 5.1: Total area in hectares (ha) and percentage of national area by Level 1 group for NLC 2018 and CLC 2018.

Chart 5.1 and Table 5.1 above, highlights that the Grasslands, Swamp and Saltmarsh class is by far the most dominant class group in the NLC 2018 data with 3.8 million ha or 54.2% of the total national area. It is also the most dominant class for CORINE 2018, but NLC 2018 reports 693,534 ha or 9.81% less grassland nationally for the same year. The Forest, Woodland and Scrub group has the second highest area in the NLC 2018 data at 1.29 million ha or 18.27% of the national area. This is an increase of 618,626 ha or 8.76% of the national area compared to CORINE. In CORINE, the Peatland group is the second most widespread land cover type at 13.71%, but this group is the third most widespread group in the NLC 2018 dataset and has an area of 462,292 ha or 6.55% of the total national area. The heath and Bracken group saw the second highest gain compared to CORINE with an increase of 4.68% to give an area of 456,917 ha or 6.47% of the total national area. In the remaining Level 1 groups, Artificial Surfaces and Cultivated Land saw an increase of just over 1.5% compared to CORINE and the Exposed Surfaces and Waterbodies saw a difference of < 1% to the CORINE data.

5.2 Level 2 Results

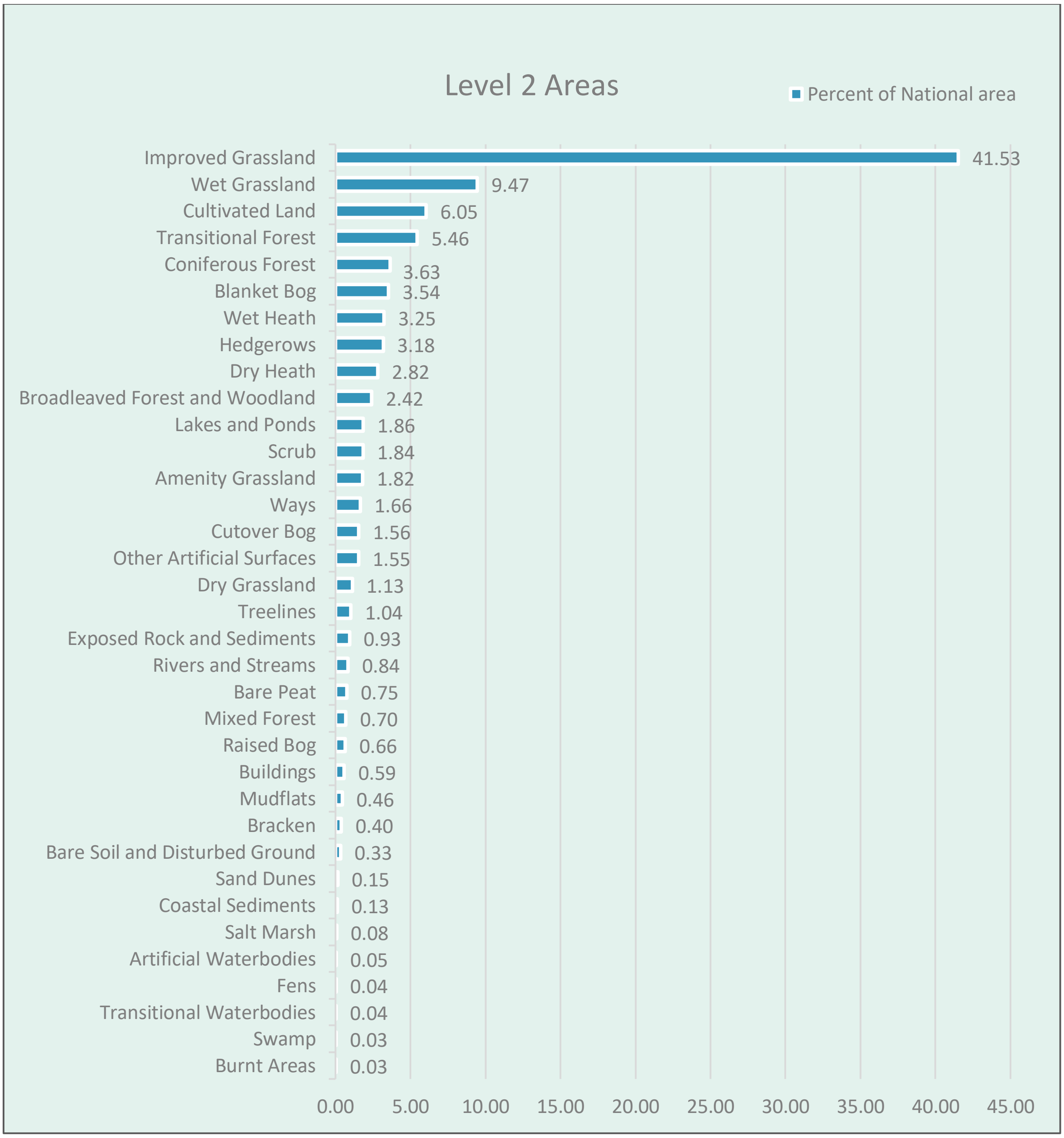


Chart 5.2: NLC 2018 Level 2 land cover classes shown in descending order as percentage of total national area.
**Note that the Marine Water class was excluded in calculating percentage values to avoid a distortion of the stats.

	Code	Level 1	Code	Level 2	Area Ha	% of Area
	100	ARTIFICIAL SURFACES	110	Buildings	41,675.72	0.59
			120	Ways	116,993.66	1.66
			130	Other Artificial Surfaces	109,346.69	1.55
	200	EXPOSED SURFACES	210	Exposed Rock and Sediments	65,984.07	0.93
			220	Coastal Sediments	9,241.58	0.13
			230	Mudflats	32,437.94	0.46
			240	Bare Soil and Disturbed Ground	23,656.74	0.33
			250	Burnt Areas	1,950.16	0.03
	300	CULTIVATED LAND	310	Cultivated Land	427,033.09	6.05
	400	FOREST, WOODLAND AND SCRUB	410	Coniferous Forest	256,443.08	3.63
			420	Mixed Forest	49,503.45	0.70
			430	Transitional Forest	385,672.93	5.46
			440	Broadleaved Forest and Woodland	170,859.69	2.42
			450	Scrub	130,097.97	1.84
			460	Hedgerows	224,787.22	3.18
			470	Treelines	73,392.53	1.04
	500	GRASSLAND, SWAMP AND SALTMARSH	510	Improved Grassland	2,933,229.96	41.53
			520	Amenity Grassland	128,564.37	1.82
			530	Dry Grassland	79,541.81	1.13
			540	Wet Grassland	668,548.02	9.47
			550	Saltmarsh	5,747.21	0.08
			560	Sand Dunes	10,261.66	0.15
			570	Swamp	2,267.67	0.03
	600	PEATLAND	610	Raised Bog	46,268.87	0.66
			620	Blanket Bog	249,738.18	3.54
			630	Cutover Bog	110,253.83	1.56
			640	Bare Peat	52,962.62	0.75
			650	Fens	3,068.05	0.04
	700	HEATH AND BRACKEN	710	Bracken	28,133.20	0.40
			720	Dry Heath	199,256.91	2.82
			730	Wet Heath	229,526.39	3.25
	800	WATERBODIES	810	Rivers and Streams	59,377.06	0.84
			820	Lakes and Ponds	131,263.30	1.86
			830	Artificial Waterbodies	3,654.47	0.05
			840	Transitional Waterbodies	2,504.98	0.04
			850	Marine Water	402,984.86	n/a

Table 5.2: Table showing the area in hectares (Area Ha) and percentage of total national area (% of area) for each NLC 2018 Level 2 land cover class.

**Note that the Marine Water class was excluded in calculating percentage values to avoid a distortion of the stats

The area statistics for the Level 2 classes show that by a large margin, Improved Grassland is the single most dominant land cover type in Ireland. It covers 2.93 million hectares or 41.53% of the total national area. It is the only class that exceeds 10% of the national area with the second most widespread Level 2 class being Wet Grassland at 668,548 ha or 9.47%. This shows that of the thirty-six land cover types identified in the Irish landscape, these two grassland classes together account for over 50% of the national area. Cultivated Land is the third most dominant land cover type in Ireland at 427,033 ha or 6.05% of the national area, it also has a strong regional concentration in the East and South East, as shown in section 5.4.3 below.

Transitional and Coniferous forest lands, both associated with plantation forestry account for 5.46% and 3.63% of the national area respectively. Blanket Bog and Wet Heath, both the most dominant class in their Level 1 groups, occupy 3.54% and 3.25% of the national area respectively. The first national scale mapping of hedgerows in Ireland show that they cover 224,787 ha or 3.18% of the total national area. All other classes are below 3% of the national area with full national-scale mapping achieved for the first time for many other land cover types including Dry Heath (199,257 ha / 2.82%), Broadleaved Forest and Woodland (170,860 ha / 2.42%), Scrub (130,098 ha / 1.84%), Amenity Grassland (128,564 ha / 1.82%), Cutover Bog (110,254 ha / 1.56%) and Raised Bog (46,269 ha / 0.66%).

5.3 National Distribution of Land Cover Groups

5.3.1 Artificial Surfaces

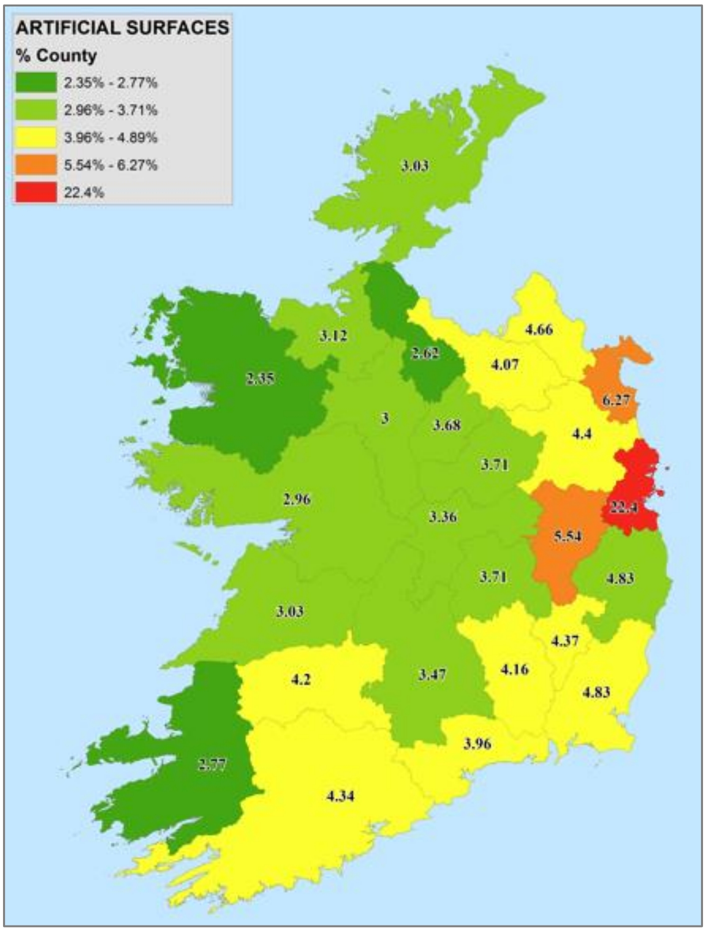


Figure 5.1: Map of Ireland showing the percentage of each county covered by the Level 1 Artificial Surfaces class.

The percentage of Ireland covered by Artificial Surfaces is 3.79%, slightly less than the EU average of 4.3% (EEA, 2020). At 22.4%, Co. Dublin has a markedly higher proportion of its area covered by Artificial Surfaces than the national median of 3.71% per county. Of the three Level 2 Artificial classes, the Buildings (110) and Ways (120) classes were mapped directly from the PRIME 2 database and account for 5.94% and 5.32% of Co. Dublin respectively. The Other Artificial Surfaces class (130) includes surfaces like driveways, farmyards, forecourts and sealed public spaces previously unmapped in PRIME2. This class accounts for 11.14% of Co. Dublin and 1.55% of the total national area (109,347 ha). The Buildings class has the highest count (i.e. number of features) of all classes with 597,281 Buildings mapped in Co. Dublin and just over 3.5 million nationally out of a total of 10.1 million features in NLC 2018.

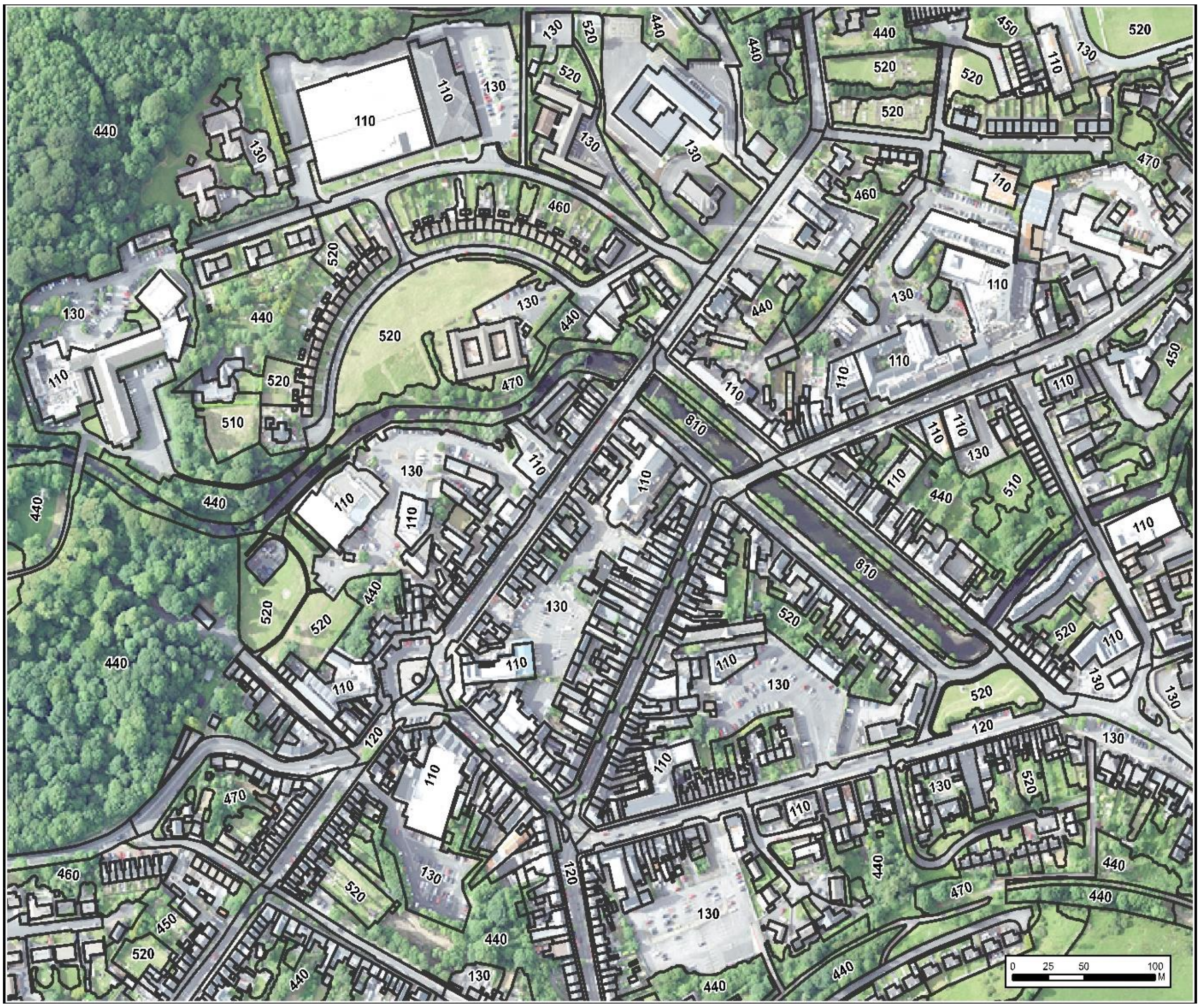


Figure 5.2: An example of the NLC 2018 map in Westport, Co. Mayo showing examples of all three Artificial Surfaces classes: Buildings (110), Ways (120) and Other Artificial Surfaces (130). Amenity Grassland (520) example can also be seen.

Figure 5.3: Examples of Artificial Surfaces classes in a non-urban setting, including Buildings (110), Ways (120) and Other Artificial Surfaces (130).

5.3.2 Exposed Surfaces

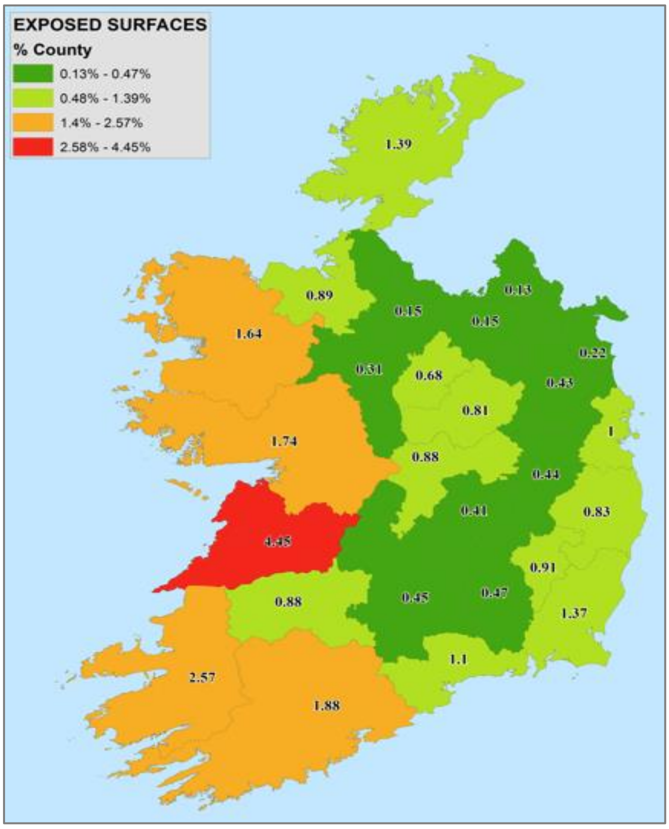


Figure 5.4: Map of Ireland showing the percentage of each county covered by the Level 1 Exposed Surfaces class.

Exposed surfaces occupy a relatively small percentage of the country at 1.89% of the national area. Over half of this is the Exposed Rock and Sediments class (0.93%) which includes exposed sections of mines and quarries, coastal rocky cliffs, wave-cut platforms, scree, and outcropping bedrock in upland areas. The distinctive limestone karst pavement in the Burren region of Co. Clare is mapped as Exposed Rock and Sediments and due to this, Co. Clare has a higher than average proportion (4.45%) mapped in the Exposed Surfaces Class. This group also includes the Coastal Sediments (220) and Mudflats (230) classes. These classes are challenging to map due to variable tidal height in the imagery with many of the mudflat areas being derived from the NPWS Coastal Monitoring Project in-situ data.

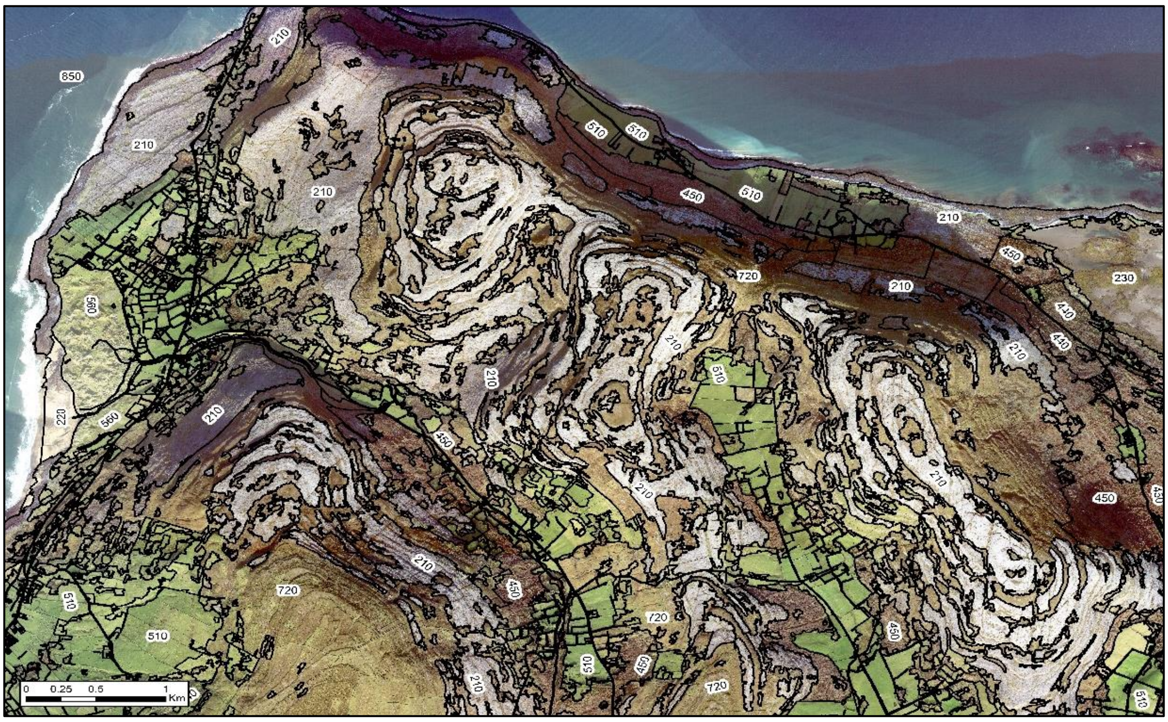


Figure 5.5: An example of the NLC 2018 map showing part of the Burren, Co. Clare. The detailed delineation of the limestone pavement can be seen, including Level 2 Exposed Rock and Sediments (210), Scrub (450) and Dry Heath (720).

Within the Level 1 Exposed Surfaces group the Level 2 classes 'Bare Soil and Disturbed Ground' and 'Burnt Areas', which were previously unmapped in Ireland, occur infrequently and are transient in nature. For this reason, these classes are challenging to map as can be seen in their respective validation results in chapter 6. However, they can be a good indicator of land cover change or environmental damage, so they are important classes to map.

5.3.3 Cultivated Land

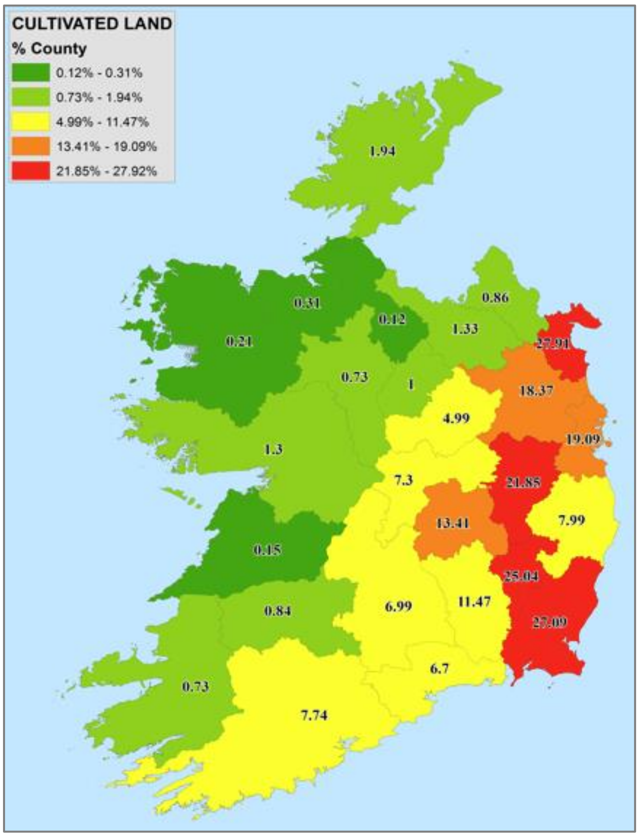


Figure 5.6: Map of Ireland showing the percentage of each county covered by the Level 1 & 2 Cultivated Land class.

The percentage of the national area mapped as Cultivated Land is 6.05% in the NLC 2018 dataset. This is higher than CORINE 2018 at 4.54%. The Cultivated Land group has a single Cultivated Land class at Level 2 which incorporates both arable and horticultural land. These agricultural land uses are identified and classified through the DAFM Land Parcel and Information System (LPIS) as described in section 4. CORINE also uses the LPIS dataset to classify its equivalent 211 class. The increase is most likely due to the more detailed spatial resolution NLC 2018. There is a strong regional trend in the location of Cultivated Land with counties in the East and South having much higher percentages of Cultivated Land than counties in the West and North West. Counties Louth, Wexford and Carlow have the highest percentage of Cultivated land at 27.91%, 27.09% and 25.04% respectively.

Figure 5.7: Examples of field parcels classified under the Cultivated Land Level 2 class.

5.3.4 Forest, Woodland and Scrub

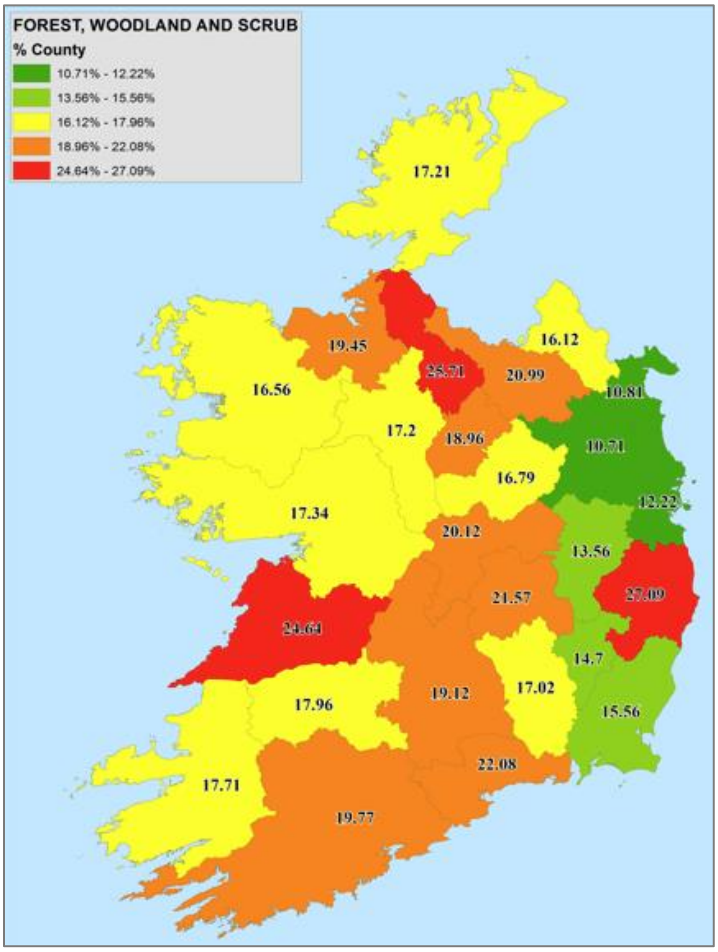


Figure 5.8: Map of Ireland showing the percentage of each county covered by the Level 1 Forest, Woodland and Scrub class.

The NLC 2018 map provides the most comprehensive spatial mapping and classification of Forest Land, Woodland and Scrub to date in Ireland. The total area of land in this NLC Level 1 class is 1.29 million ha or 18.27% of the total national area.

This figure is a significant increase from other data sources on Forest Lands in Ireland with the 2017 National Forest Inventory (NFI 2017) giving a figure of 770, 020 ha or 11% and CORINE 2018 giving a figure of 672,084 ha or 9.51%. Counties Wicklow, Leitrim and Clare have the highest percentage of forest at 27.09%, 25.71% and 24.64% respectively.

The majority of Coniferous (410), Mixed (420) and Transitional Forest (430) occur as part of a private grant-aided (PGA) or commercial forestry plantation.

As with CORINE 2018, these forest types have been mapped with the aid of data from the Forest Service, within DAFM, and Coillte. Correspondingly, combining area statistics for Coniferous (410), Mixed (420) and Transitional Forest (430) classes within NLC 2018 shows that they account for 691,619 ha or 9.79% of the national area, these figures are closer to those of CORINE 2018 and NFI 2017.

The relatively large increase of 7.28% and 8.7% seen when comparing NLC 2018 to the NFI 2017 and CORINE 2018 respectively, is then mostly in part to the detailed mapping of additional Broadleaved woodland (440), Scrub (450), Hedgerows (460) and Treelines (470) in NLC 2018. Collectively these classes comprise 8.48% of the national area. Of these four classes, Hedgerows are the most widespread accounting for 224,787 ha or 3.18% of the total national area.

It is important to note that not all the Level 2 classes in the Forest, Broadleaved Woodland and Scrub Level 1 class would necessarily be considered as 'Forest' in some sectors. This means that the different forest classifications need to be accounted for when comparing figures collated for different purposes. It is also worth noting that Level 2 classes within the Level 1 Forest, Broadleaved Woodland and Scrub class, are defined by their forest type and spectral properties rather than their land use, state-of-naturalness, or habitat type. Examples of some of the different Level 2 Forest classes can be seen in Figure 5.9 – 5.10 below.

Figure 5.9: Example of the NLC 2018 map showing areas of the Coniferous Forest (410), Mixed Forest (420), Transitional Forest (430) and Broadleaved Forest (440).

Figure 5.10: Example of the NLC 2018 map showing areas of the Broadleaved Forest (440), Scrub (450), Hedgerows (460) and Treeline (470) classes.

5.3.5 Grassland, Swamp and Saltmarsh

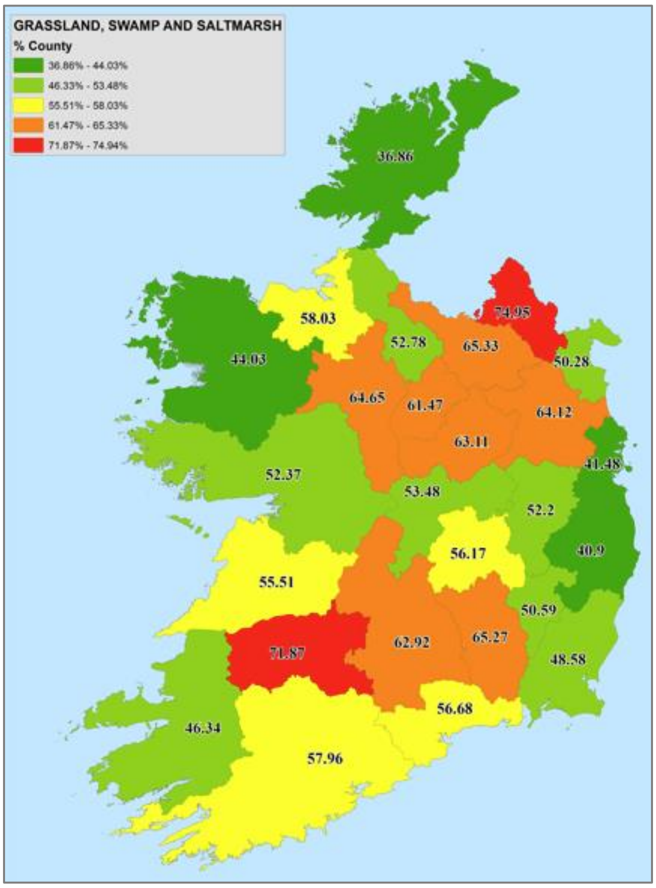


Figure 5.11: Map of Ireland showing the percentage of each county covered by the Level 1 Grassland, Swamp and Saltmarsh Level 1.

The Grassland, Swamp and Saltmarsh Level 1 class is the most widespread land cover class in NLC 2018 at 3.8 million ha or 54.20%. Counties Monaghan and Limerick have the highest proportion of Grassland at 74.95% and 71.87% respectively. Although the land use of these grasslands is not specified, this indicates that both counties would predominantly be covered in agricultural grasslands. NLC 2018 provides the first national scale mapping of Amenity Grassland (520) covering 128,564 ha or 1.82% of the national area. Also, the first national scale mapping of wet and dry grasslands was undertaken. Although chapter 6 shows that these two classes were particularly challenging to map consistently, with lower accuracy results relative to other classes. Mapping of Sand Dunes, Swamp and Saltmarsh classes also provides national-scale spatial information for these semi-natural land cover types which are important for local scale biodiversity management.

Figure 5.12: Example of the NLC 2018 map showing areas of the Improved (510) and Wet (540) Grassland classes.

Figure 5.13: Example of the NLC 2018 map around Doolin, Co. Clare showing areas of Dry Grassland (530), Improved Grassland (510), Amenity Grassland (520) and Sand Dunes (560)

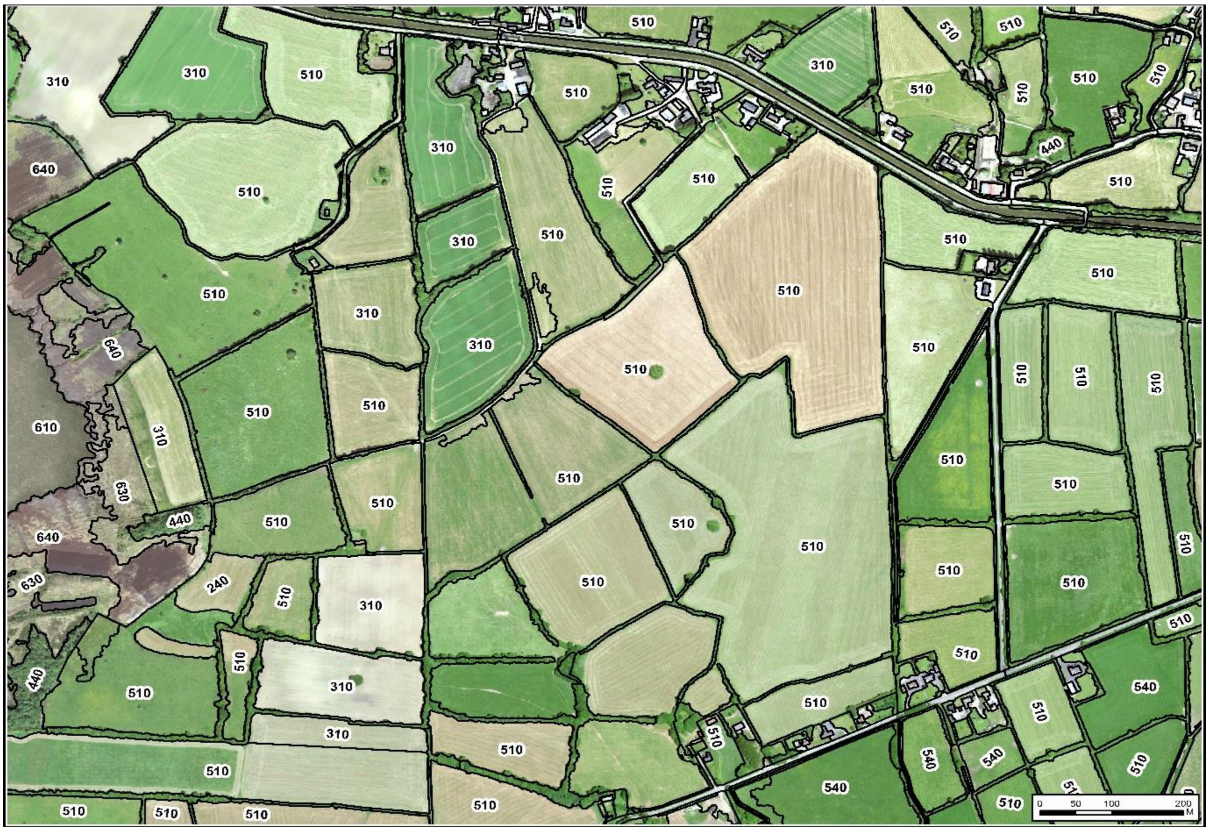


Figure 5.14: Example of the NLC 2018 from Co. Offaly showing different stages of Improved Grassland (510) during the summer cutting season.

5.3.6 Peatland

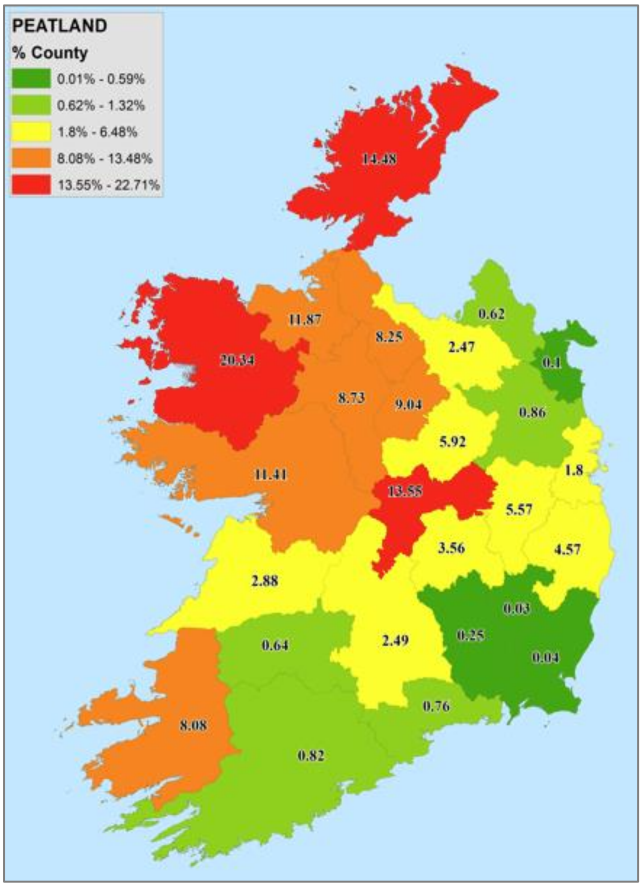


Figure 5.15: Map of Ireland showing the percentage of each county covered by the Level 1 Peatland class.

The NLC 2018 Peatland Level 1 class includes five Level 2 classes covering the main peat surfaces found in Ireland: Raised Bog (0.66%), Blanket Bog (3.54%), Cutover Bog (1.56%), Bare Peat (0.75%) and Fens (0.04%). In total the Level 1 Peatland class covers 6.55% of national area which is considerably less than CORINE at 13.7%. The main reason for the difference is the lower 25 ha resolution of CORINE, which does not identify and map different land cover types that occur within or on a peat complex. It also generalises small features into the Peatland class. At 6.55%, the amount of peatland identified in NLC 2018 is also well below the figure of 20% of national area given by the Derived Irish Peatland Map (DIPM) (Connolly & Holden, 2009). However, a follow-on study which mapped the land uses on these peat complexes within DIPM found that 66% of these had either a grassland, forestry or industrial land use, leaving just 35% ‘residual peat’ (Connolly, 2018) which is equivalent to 6.8% of the national area.

It is important to note that there are different interpretations of what constitutes a peatland and that this land cover type has seen large scale modification over time and in the landscape due to the uses of peatlands. Along with peat harvesting, many areas of peatland have been drained and either afforested with plantation forestry or converted to grassland. There is also evidence of encroachment of other land cover types in areas that are not actively managed. A peat ‘complex’ can therefore have many different land cover types present on-top of a peat soil.

NLC 2018 maps peat surfaces only, it does not map peat soils or label land cover which occurs on peat soils as e.g. ‘forest on peatland’. There are other projects, like the DAFM / EPA funded REPEAT Research Project in Trinity College Dublin that aims to identify organic peat soils. Figure 5.16 below provides a good illustration of the different land cover types including Scrub, Transitional Forest and Lakes and Ponds, that can occur on a modified peat complex.

Figure 5.16: Example of different land cover types found on an industrially exploited raised bogs in the NLC 2018 map.

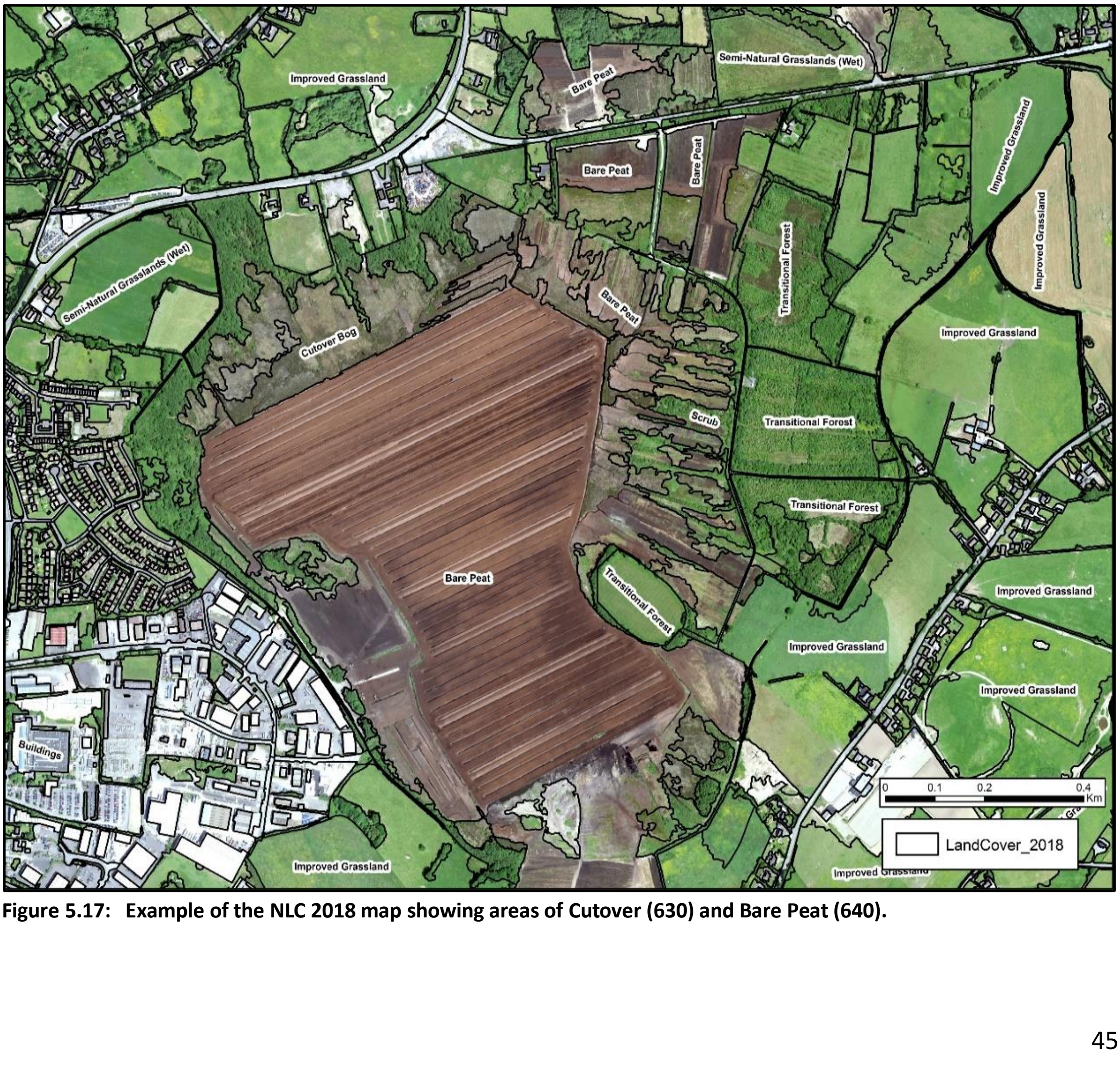


Figure 5.17: Example of the NLC 2018 map showing areas of Cutover (630) and Bare Peat (640).

Figure 5.18: Example of Blanket Bog (620) and Cutover (630) at Lough Keel, Achill Island, Co. Mayo.

5.3.7 Heath and Bracken

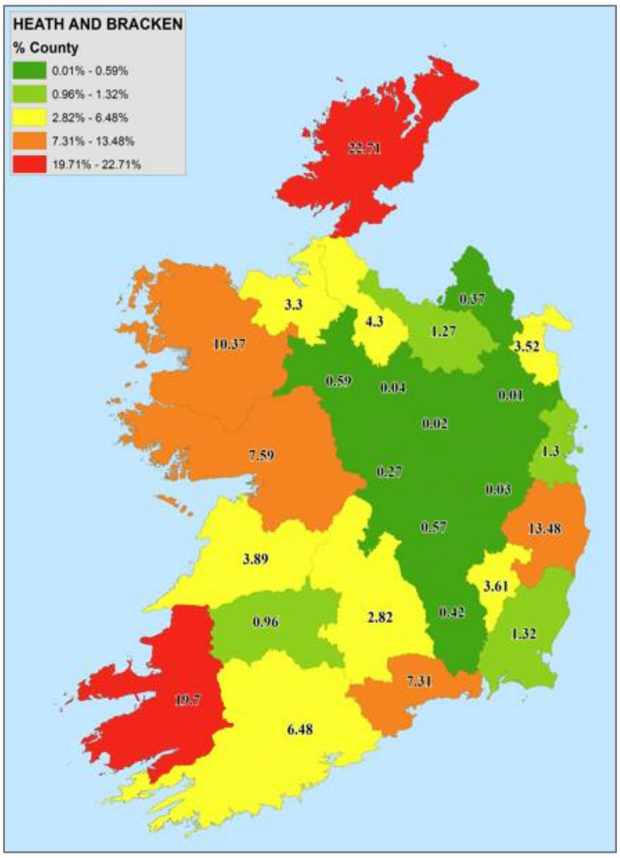


Figure 5.19: Map of Ireland showing the percentage of each county covered by the Level 1 Heath and Bracken class.

The Level 1 Heath and Bracken class is the fourth most widespread land cover group in NLC 2018 at 6.47% of the national area. This group separates heath land into Wet Heath (3.25%) and Dry Heath (2.82%), which are not distinguished in CORINE where one Heathland class represents 1.78% of the national area.

Wet Heath can occur frequently in upland habitats where it is mosaiced with both peatland and wet grassland classes. Considering this the difference between NLC 2018 and CORINE is most likely due to areas of Wet Heath being classified as Peatland in the CORINE dataset. NLC 2018 shows that Donegal and Kerry have significant areas of Heath and Bracken with 22.71% and 19.70% of their respective county areas. NLC 2018 also provides the first national scale mapping of Bracken, a pernicious species which rapidly colonises free-draining soils in upland and coastal areas and on the margins of agricultural land and waste ground.

Figure 5.20: Example of the NLC 2018 map showing areas of Wet Heath, Dry Heath, and Blanket bog in Connemara.

Figure 5.21: Example of areas of Bracken, Dry Heath, Exposed Rock and Sediments, and Scrub in Co. Wicklow.

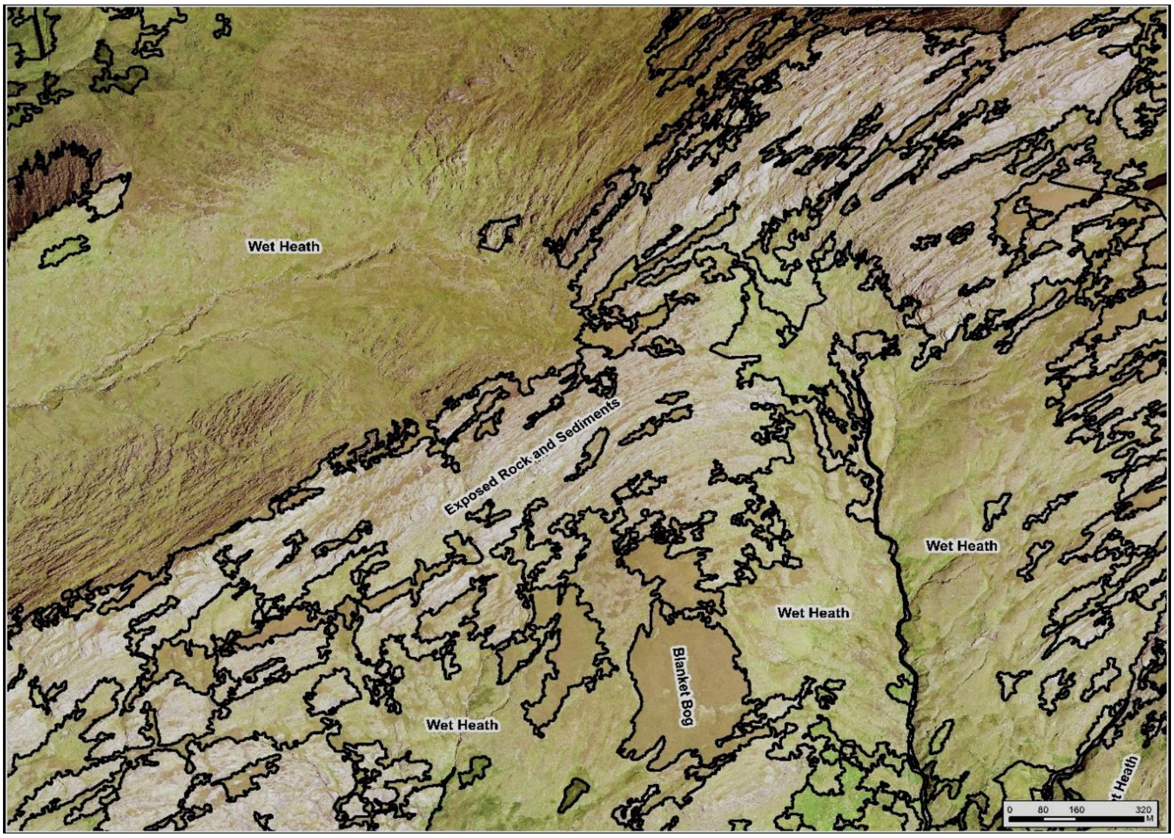


Figure 5.22: Example of the NLC 2018 map showing areas of Wet Heath, Blanket Bog, Bracken and Dry Heath in Co. Cork.

5.3.8 Waterbodies

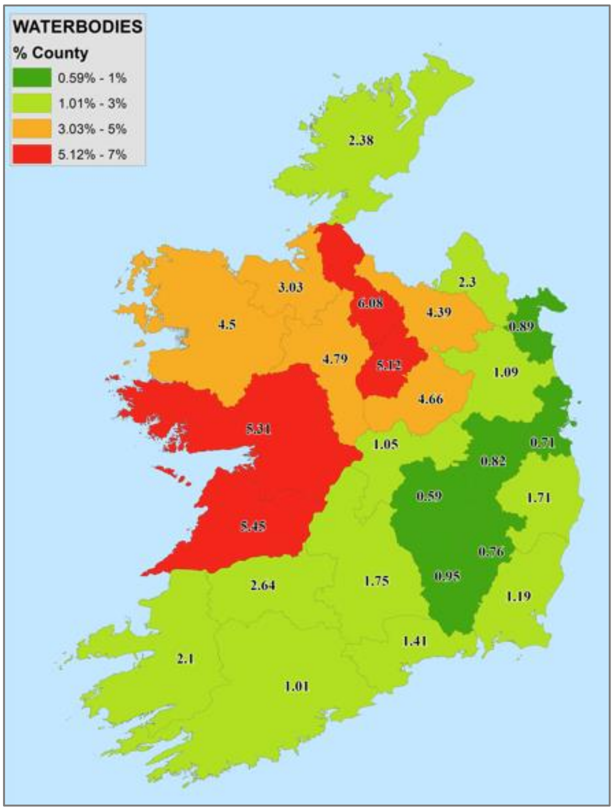


Figure 5.23: Map of Ireland showing the percentage of each county covered by the Level 1 Waterbodies class.

The Waterbodies Level 1 class includes Level 2 classes that were primarily taken from the PRIME2 dataset, meaning that the area and distribution of waterbodies will match this dataset closely. NLC 2018 shows that the Level 1 Waterbodies represent 2.79% of the national area, at the Level 2 classes this consists of Rivers and Streams (0.84%), Lakes and Ponds (1.86%), Artificial Waterbodies (0.05%) and Transitional Waterbodies (0.04%). There is a strong regional variation in the proportion of each county covered in waterbodies. Inland counties in the South east such as Counties Laois, Dublin, Carlow, Kildare, Louth, and Kilkenny have <1% of the county covered in waterbodies. Counties Clare, Galway, and Longford > 5% of their area as water. While Co. Leitrim shows the highest proportion of area as water at 6.08%.

The Marine Water class is not included in the calculation of national area statistics.

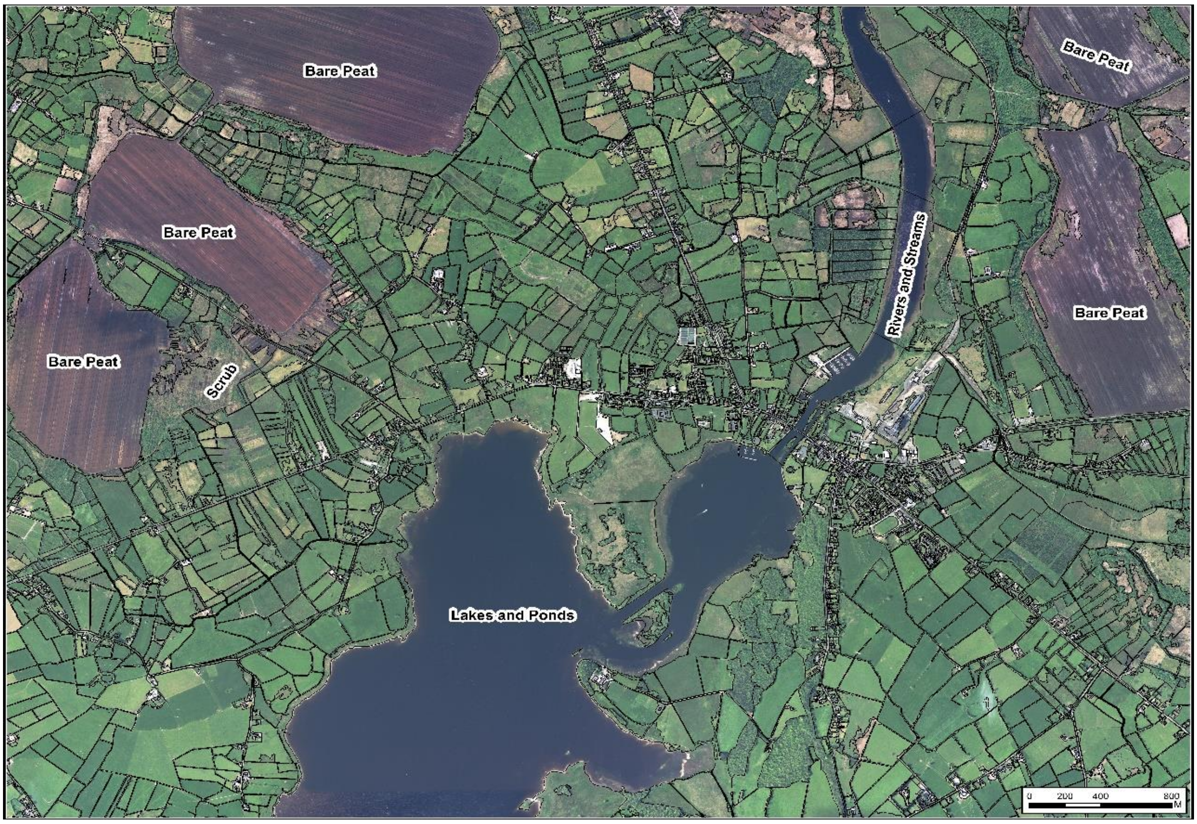


Figure 5.24: Example of the NLC 2018 map showing the Shannon entering Lough Ree in Co. Longford & Co. Roscommon.

6 External Validation & Accuracy Assessment

A key element of the NLC 2018 was the design and implementation of an extensive data validation programme which ran in parallel to the main production process. Its purpose was to provide independently assesed and verified statistics on the accuracy of the NLC 2018 data and to give end users a deeper understanding and confidence in the final product. The validation programme was designed by the EPA, OSi and Central Statistics Office (CSO) in consultation with stakeholders and in accordance with spatial mapping, remote sensing, and statistical best practices.

Essential to the programmes success was the participation of independent validation partners. The EPA and OSi wish to thank and recognise the considerable contributions of staff from the following organisations who helped to manually validate over 22,000 samples:

- BEC Consultants Ltd.
- Central Statistics Office
- Department of Agriculture, Food & the Marine
- Atlantic Technological University
- National Parks & Wildlife Service
- Teagasc
- Environmental Protection Agency

This section provides an overview of the validation programme and accuracy assessment results. A more in-depth validation and accuracy assessment report will follow after the data publication.

6.1 Validation Overview

As extensive field-based, ground truthing of the data product was not feasible, a extensive product validation was undertaken of to quantify and improve the quality of NLC 2018. Product validation is an assessment of how accurately the map has been produced in reference to the core production data, not how well it reflects/compares with the reality on the ground (ground-truthing).

Product validation was undertaken using manual photo-interpretation, where the validator was issued with a set of samples from the final NLC 2018 dataset, which, they then compare visually against the OSi's Series 2 aerial ortho photography images and other core raster and vector data sets.

Validation was undertaken by a team of external validation partners composed of a total of 54 domain experts in ecology and habitat mapping, GIS & Remote Sensing, land use and agricultural land management. Extensive online training was provided to all validators and a web-application developed to provide a standardised platform to undertake the validation work.

In parallel with the map production process, validation was stuctured across Biogeographical Regions (BGRs). Once all working blocks in a BGR were classified, they were stitched together and validated as a single dataset, with the sample selection and stratification undertaken across BGRs. Consequently, validation metrics were calculated at both BGR and national aggregation levels.

In total, 21,614 samples were validated by the independent validation partners. The project teams' aim was to produce a map with an overall accuracy >75%.

6.2 Product Validation

The product validation approach for NLC 2018 was a scientifcally rigorous, combined qualitative and quantitative assessment.

6.2.1 Quantitative Validation

The quantitative validation process was the core statistical assesment of the data product. It involved the validation of a set of stratified random samples from the NLC 2018 data product for each BGR. All validators were assigned a unique and random subset of the entire sample set to validate. They were asked to visually assess each land cover sample for thematic and geometric accuracy (i.e. is the classification label and spatial boundary accurate), this was done using the tools and data provided in a dedicated web application.

In addition to the random samples a subset of 1,652 'common samples', representing 7.4% of the total sample set, were distributed to all validation partners. Common samples were not identifiable to validators to ensure consistent validation across both random and common samples, and facilitated cross-validaiton of validation performance across partners. The project teams' ecological service partners (BEC Consultants Ltd.), were tasked to manually assess and classify all common samples with the aim of establishing the 'most likely land cover' class. For each common sample the validation response of the validation partners was compared directly against BEC's 'most likely land cover' classin order to assess validation quality, consistency and reliability across validators, classes and regions.

6.2.2 Qualitative Validation – 'Look and Feel' Assessment

In addition to quanitative validation, validation partners conducted a qualitative, 'look-and-feel' assessment. This was a more unstructured and holistic assesment, where validators were asked to visually assess the data, evaluating the overall quality and feel of the data product, checking for systematic and/or isolated errors that may not have been picked up in samples in the quantitiative assesment. Validators provided feedback from this assessment in each of their BGR validation reports, with screenshots and descriptions of commonly identified issues.

6.3 Metrics Used in the Assessment of Land Cover Data Accuracy

There are a number of metrics that can be applied to measuring the accuracy of a classified map. The key metrics used for assessing the accuracy of NLC 2018 are explained below.

6.3.1 Predicted vs. Actual Value

When assesing classification accuracy, it is important to compare the predicted value verses the actual value. In the case of NLC 2018, the predicted value is the classification label given to a polygon by the classifier and the actual value is the label manually assigned to the polygon by the validator. A direct comparison of these predicted vs. actual values can provide a simple 'overall' accuracy result. However, in the analysis of remote sensing data, error matrices are often used where all predicted values are plotted against all the acutal values which allows more in-depth analysis of classification accuracy across classes including identifying misclassification trends between classes. Figure 6.1 below shows a simplified confusion matrix and the different possible relationships between the Predicted and Actual Class values. A detailed confusion matrix for all classes in the land cover validation process is provided in Appendix 4.

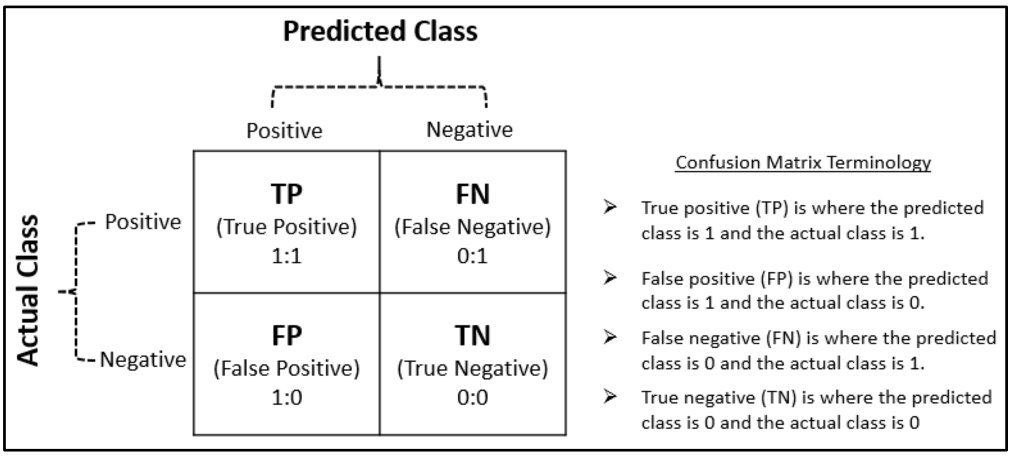


Figure 6.1: Diagram explaining confusion matrix and associated terminology.

6.3.2 Accuracy Metrics

A range of accuracy metrics, based on an assesment of the predicted vs actual validation class, were used to generate detailed accuracy metrics on the NLC 2018 Map. These are detailed in table 6.1 below:

Producer's Accuracy	Calculates from all the positive samples in a class (TP + FN), how many were predicted correctly in the map (TP). Also known as "Observed Accuracy", "Recall" or "1 - Omission Error" Calculation: TP/ (TP + FP)
User's Accuracy	Calculates from all the samples in a class predicted as positive in the map (TP + FP), how many are actually positive (TP). Also known as "Precision" or "1 - Commission Error" Calculation: TP/ (TP + FN)
F1 Score	Combines the producer's and user's accuracy into a single balanced metric. Calculation: 2 * (Users Accuracy * Producers Accuracy) / (Users Accuracy + Producers Accuracy)
Kappa	Summarises the degree of agreement between two or more sources via the ratio of the observed proportion agreement over the expected proportation agreement by chance. It is a metric best used when comparing maps across different producers.

Table 6.1: Table showing the main metrics used in the land cover validation process

It is worth noting that the confidence intervals reported with accuracy metrics below are calculated using bootstrapping and fairly reflect the uncertainty around a metric based on the underlying variation in the data being aggregated to calculate the metric. .

6.4 Accuracy Results

The selected accuracy metrics were then calculated to assess the quality of both the thematic classes and their associated geometry. The initial analysis for this report was applied on a national basis using Levels 1 and 2 of the classification; the results of the random sample analysis is shown in Table 6.2 below.

Accuracy Metric	**Class**	
	Level 1	**Level 2**
Producer's accuracy (%)	88.8 ± 3.0	78.5 ± 4.7
User's accuracy (%)	88.8 ± 5.6	78.5 ± 4.1
F1 score (%)	88.7 ± 3.7	78.5 ± 3.9
Kappa (%)	86.2 ± 3.4	77.4 ± 3.6
Geometric accuracy (%)	87.2 ± 22.5	87.2 ± 22.5

Table 6.2: Full breakdown of the overall NLC2018 accuracy results at Level 1 & 2.

A future in-depth validation accuracy assessment report will provide detailed assessments on variances between biogeographical areas and across the validation partners.

6.4.1 Overall Accuracy

The overall accuracy of the NLC 2018 at Level 1 and 2 of the classification is quite consistent across the four metrics used. Using the weighted mean of each metric across all land cover classes results in an overall accuracy of between 86 - 89% at Level 1 and 77 - 79% at Level 2.

The geometric accuracy is 87% for both levels, this is to be expected as the geometry was mapped and validated at Level 2 with no geometry changes when aggregating to Level 1 classes.

6.4.2 Producers and Users Accuracy

Combined producer's and user's accuracy results for individiual classes at Level 1 and 2 are provided in Figures 6.2 and 6.3 below. This analysis shows that overall there is good classification performance across Levels 1 and 2, but that some Level 2 classes drop below the >75% threshold. This is to be expected as the Level 1 (aggregated) classes are more general land class groupings and are more distinct from each other compared to Level 2 classes. Figure 6.2 shows that at Level 1 the map is performing strongly with all accuracy values >75% and most >80%. The artificial, forest, grassland and waterbody classes show the best results with exposed surfaces and heath and bracken classes performing less well.

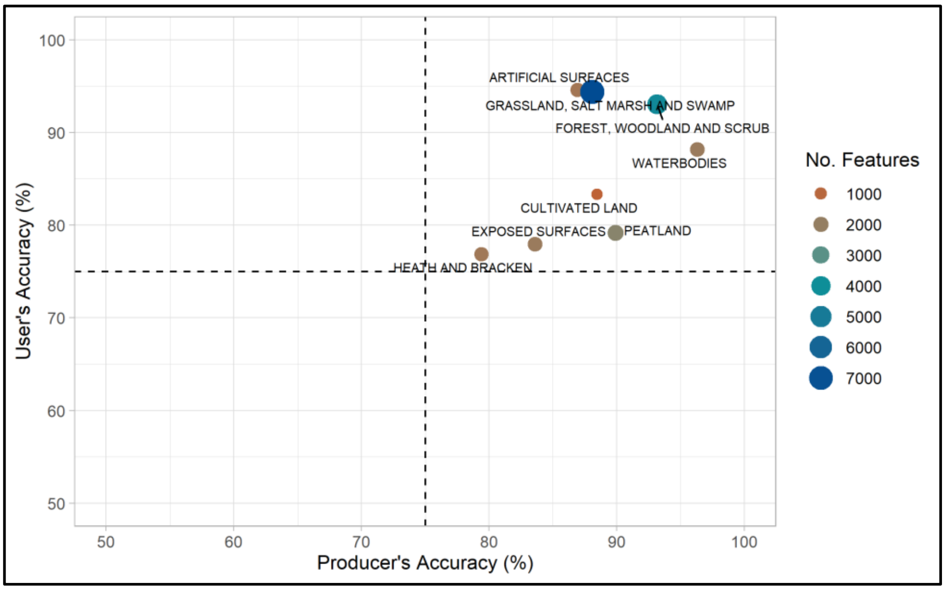


Figure 6.2: NLC 2018 Level 1 producer's vs. user's accuracy assessment.

Similarly, Figure 6.3 demonstrates that at Level 2 the map is also performing well overall, with 15 classes achieving >75% in both producer's and user's accruacy and 10 classes >80%. Some of the best performing classes are Buildings, Ways, Cultivated Land, Hedgerows, Improved Grassland and Saltmarsh. However, there are accuracy issues with some classes being <75% and some even <60%. More challenging classes performed less well such as Bare Soil, Burnt Areas, Dry Grassland, Wet Grassland, Scrub and Cutover Bog.

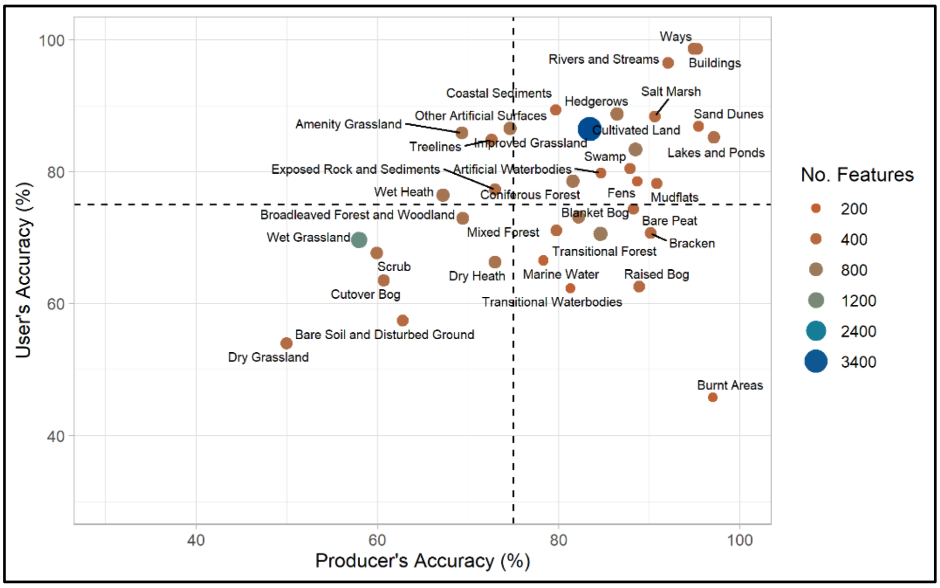


Figure 6.3: NLC 2018 Level 2 producer's vs. user's accuracy assessment.

6.4.3 F1 Score and Kappa

The F1 Score, also known as the Dice coeffecient, combines both the producer's and user's accuracy and more fairly reflects the overall accuracy of a class. Figure 6.4 below provides an F1 Score analysis of all Level 2 classes. Overall, 24 classes had >75% accuracy and 17 classes >80%. There are still some accuracy issues with 12 classes <75% and one class (Dry Grassland) has >60% in accuracy. The Kappa coefficient expressed as a percentage shows very similar results across classes to F1 Score (Figure 6.4).

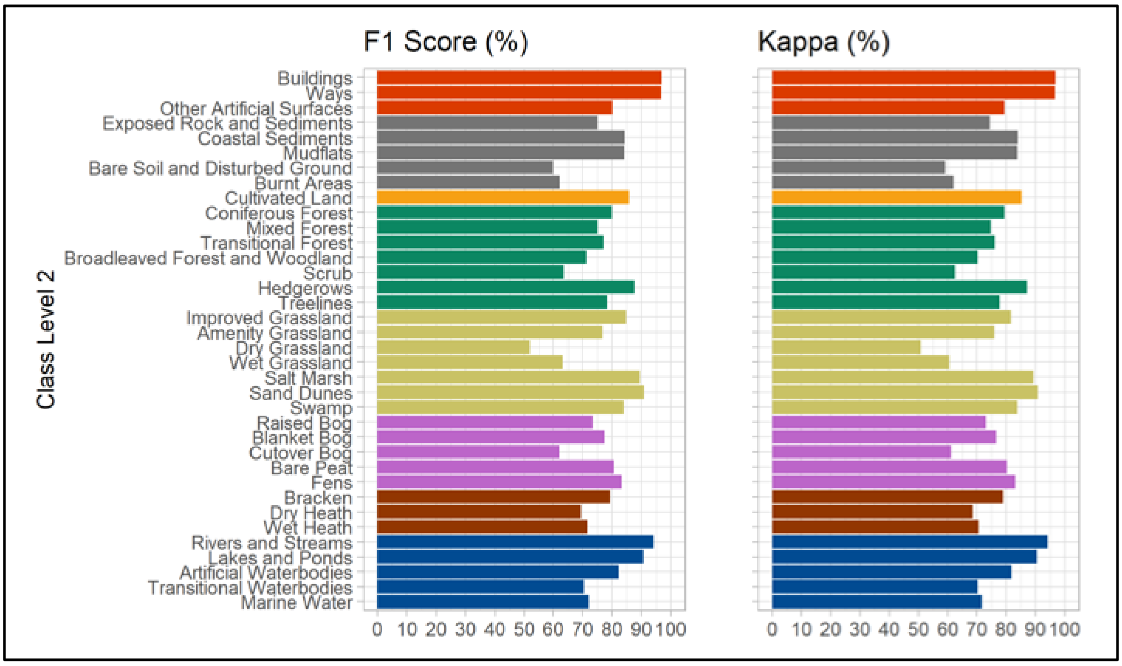


Figure 6.4: F1 Score and Kappa accuracy assessment results per NLC 2018 Level 2 class

6.4.4 Geometric Accuracy

During the validation process validators assessesed the geometric accuracy of the samples, scoring how well the spatial boundaries of the sample object match the ground feature being mapped within the source image. A range of accuracy bands were given in 10% increments from 0 - 100%.

The overall geometric accuracy for each Level 2 class was then calculated using the mean of all validation scores, weighted by the total no. of features in a class. This gives a more appropriate measure of central tendency when differences in the number of features across classes is relatively large. The weighted mean geometric accuracy of Level 1 & 2 samples shows accurate delineation of objects was achieved, with all Level 2 classes scoring >80%, the majority >85% and some classes >90%.

Figure 6.5 summarises the geometric accuracy across Level 1 land cover classes. All but one class, Forest, Woodland and Scrub (84%), were >85% accuracy with waterbodies, artificial surfaces and cultivated land performing particularly well. The confidence intervals highlight that there can be high variability within classes, this means that users should be aware that there can be incidence of significant geometric error.

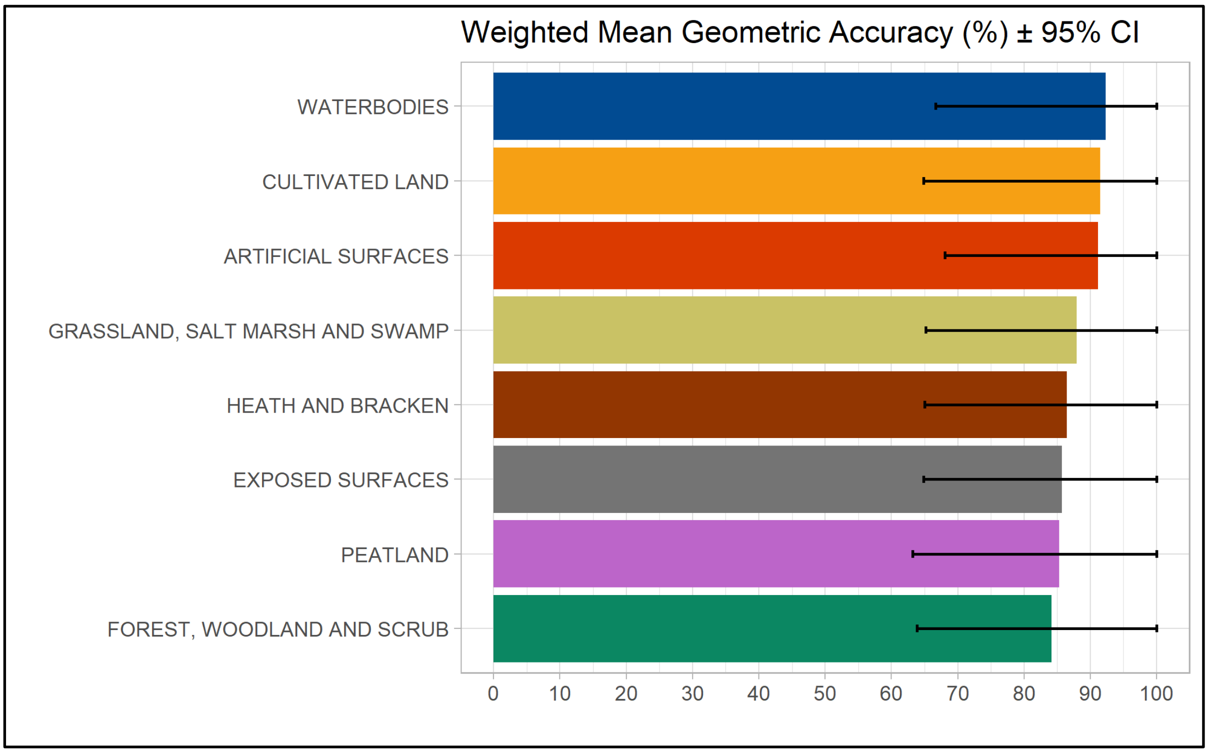


Figure 6.5: National geometric accuracy assessment of NLC 2018 Level 1 classes

All Level 2 classes were >80% in geometric accuracy. Figure 6.6 below shows the best performing classes were ways, buildings, water classes, cultivated land and improved grasslands while marine waters, scrub, cutover bog and treelines performed slighlty less well. Similar to Level 1, the confidence intervals highlight a high variability of geometric accuracy within classes, particularly in the case of marine water and cutover bog. Again, this means that users should be aware that there can be incidence of significant geometric error within classes.

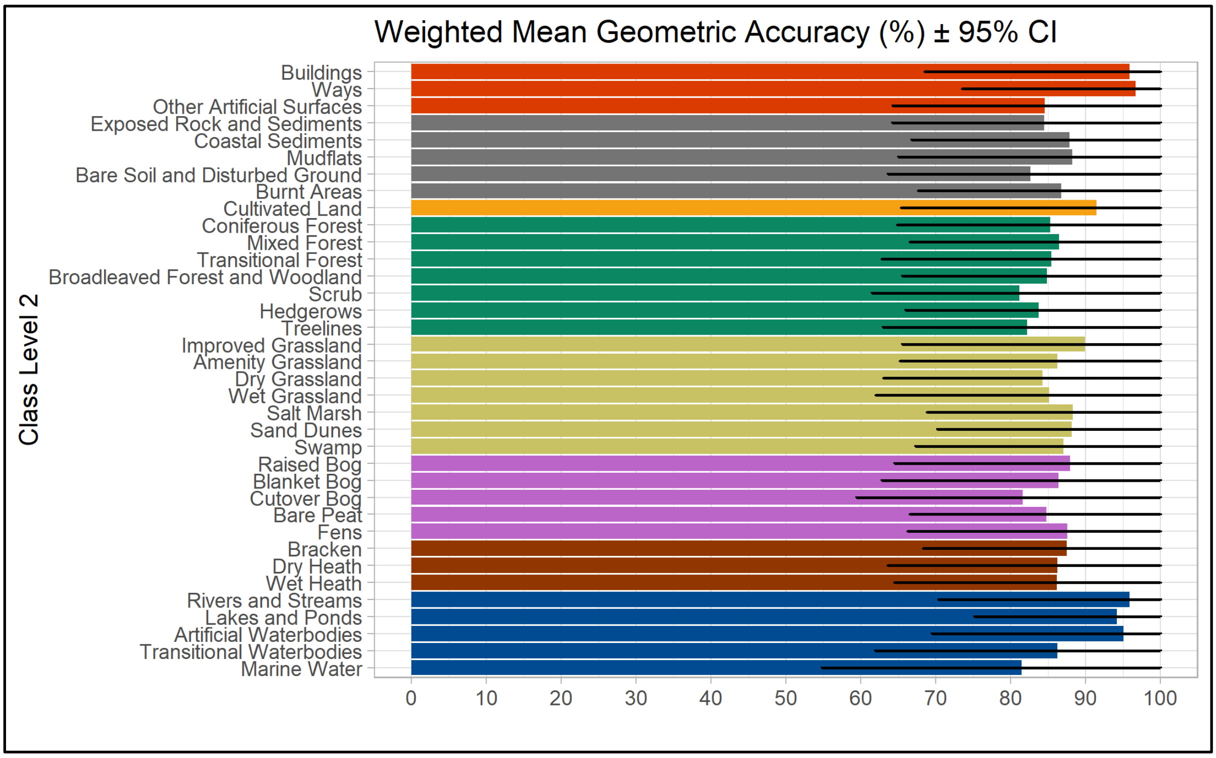


Figure 6.6: National geometric accuracy assessment of NLC 2018 Level 2 classes

6.5 Comparison of Validation Partners

As detailed in 6.1 and 6.2, validation partners were required to validate a sub set of common sample points, the purpose being to identify patterns and validation quality across classes, partners and regions. With one validation partner, BEC Consultants Ltd, providing additional resources to establish a 'most likely land cover' class, all other validation partners were then compared against this standard to facilitate internal quality control of validation performance. The results of this assessment were very positive with all validation partners performing to a high standard. Figure 6.7 below shows the performance, for Level 2 validation assessments, across validation partners for the primary metrics used in the accuracy assessment process.

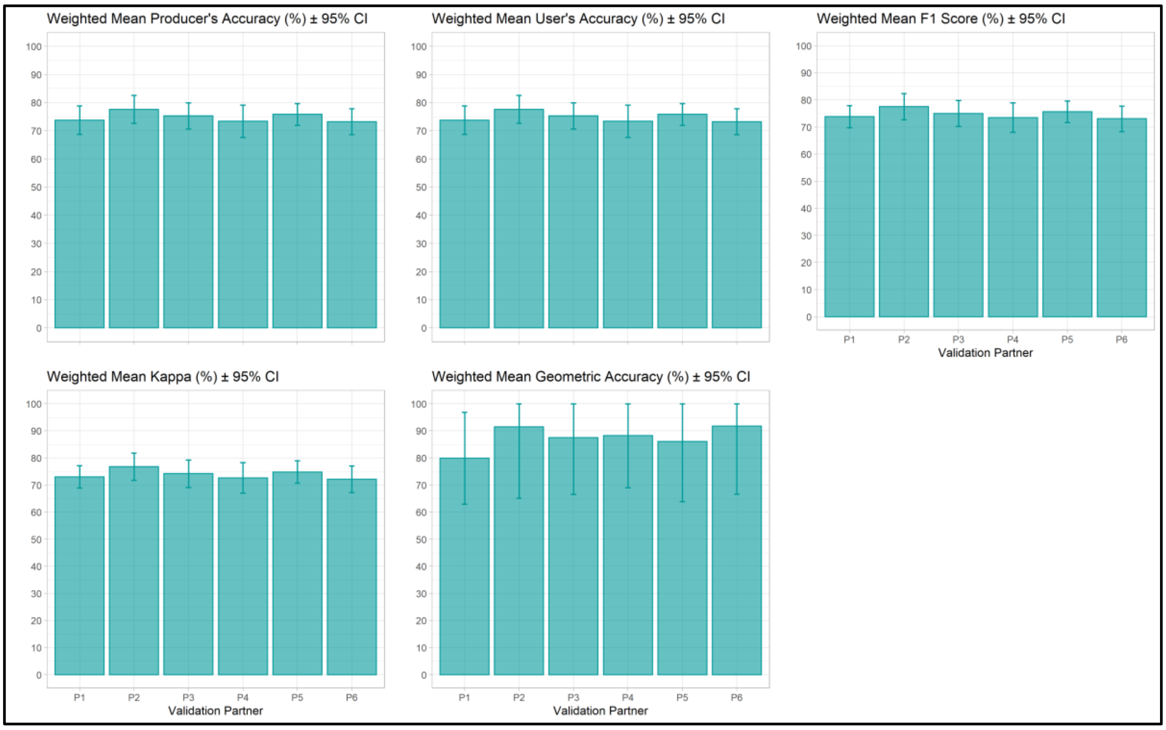


Figure 6.7: Level 2 national summary comparison across validation partners

The above metrics show that the validation partners all performed the validation task at a consitently high and similar standard. Geometric accuracy was the metric with the most variablity which aligns with feedback from validators that this task was the least clear and most subjective within the validation process.

Overall, this comparison highlights the consistency of knowledge and understanding within the validation partners in addition to the quality of the training, documentation and support provided. The consistency across the metrics also provides a high level of confidence in the validation programme and its accuracy assessment.

7 Conclusions

The National Land Cover Map of Ireland, NLC 2018, is now complete and available via the Ordnance Survey of Ireland. It is the culmination of many years of collaborate work between the OSI, the Environmental Protection Agency and the National Land Cover and Habitats Mapping working group. Through this collaboration it has enabled the design of a National Land Cover Classification System, the design and implementation of a mapping methodology, the design of a validation process and the implementation of this through independent validation partners.

NCL 2018 has been designed and developed using best practice and continuous engagement with stakeholders, with the aim of delivering the best possible product for all. The validation results show that NLC 2018 is of a very high overall quality and accuracy, it allows users to identify which land cover classes are best and which perform less well, this will help stakeholders use the data in confidence by factoring in performance results. It will also allow for more targeted improvements into the future.

The data presents new baseline information on the status of land cover and the environment in Ireland. It is hoped that this will help inform and enable more accurate and reliable assessment, reporting, research, and monitoring of the environment, while helping Ireland to tackle the many challenges presented by the Climate and Biodiversity emergencies. It is envisaged that this will be the first in a long-term series of national land cover data, which through time will help us identify and manage the changes and pressures on Ireland's Environment.

This report aims to highlight how the map was developed, what it looks like, how accurate it is and summarise what it means in terms of a new baseline of national land cover statistics. It is envisaged that the report will be followed by more detailed papers on the classification, methodology and validation process and results.

References:

- Arnold, S., Kosztra, B., Banko, G., Smith, G., Hazeu, G., Bock, M. and Valcarcel Sanz, N., 2020. The EAGLE concept—A vision of a future European Land Monitoring Framework. In *Proceedings 33th EARSeL Symposium towards Horizon* (Vol. 2020, pp. 551-568).
- Banko, G., Mansberger, R., Gallaun, H., Grillmayer, R., Prüller, R., Riedl, M., Stemberger, W., Steinnocher, K. and Walli, A., 2014. Land Information System Austria (LISA). In Land Use and Land Cover Mapping in Europe (pp. 237-254). Springer, Dordrecht.
- Ben-Asher Z., 2013 - HELM-Harmonised European Land Monitoring: Findings and Recommendations of the FP7 HELM Project. Tel-Aviv, Israel, 80 p.
- Blaschke, T., Lang, S., Hay, G. and Lang, S., (2008) Object-based image analysis for remote sensing applications: modelling reality – dealing with complexity, Object-Based Image Analysis. Lecture Notes in Geoinformation and Cartography, Springer Berlin Heidelberg, pp. 3-27
- Büttner, G., Kosztra, B., Maucha, G., Pataki, R., Kleeschulte, S., Hazeu, G.W., Vittek, M., Schroder, C. and Littkopf, A., 2021. *Copernicus Land Monitoring Service-CORINE Land Cover. User Manual*. Copernicus Publications.
- Cawkwell, F., Raab, C., Barrett, B., Green, S. and Finn, J. (2016) Toward Land cover Accounting and Monitoring (TaLAM). STRIVE Project report, Environmental Protection Agency
- Connolly, J. and Holden, N.M., 2009. Mapping peat soils in Ireland: updating the derived Irish peat map. *Irish Geography*, *42*(3), pp.343-352.
- Connolly, J., 2018. Mapping land use on Irish peatlands using medium resolution satellite imagery. *Irish Geography*, *51*(2), pp.187-204.
- Di Gregorio, A., 2005. Land cover classification system: classification concepts and user manual: LCCS (Vol. 2). Food & Agriculture Org. EEA, 2020. The European environment — State and outlook 2020. Kongens Nytorv 6, 1050 Copenhagen, Denmark
- EEA, 2021 CORINE Land Cover User Manual. Buttner, G., Kosztra, B., Kleeschulte S., Hazeu G.,
- EPA, 2020. Ireland's Environment 2020 - An Assessment. Johnstown Castle Estate, Co. Wexford, Ireland
- Fossitt, J., (2000) A guide to habitats in Ireland. The Heritage Council of Ireland Series, Heritage Council
- Green, S., and Fealy, R. 1995 Teagasc-EPA Soils and Subsoils mapping project: Final report V.1
- Green, S. and Fealy, R., 1995. Teagasc 1995 Land Cover Map of Ireland.
- Jackson, D.L. 2000. Guidance on the interpretation of the Biodiversity Broad Habitat Classification (terrestrial and freshwater types): Definitions and the relationship with other classifications, JNCC Report No. 307, JNCC, Peterborough, ISSN 0963-8091.
- Lydon, K., EPA 2016. Desk Study on Proposal for a new national Land Cover Classification System for Ireland. Unpublished.
- Kosztra, B. & Büttner G. 2019 Updated CLC illustrated nomenclature guidelines. European Topic Centre on Urban, land and soil systems, Environment Agency Austria; EAA Spittelauer Lände 5 1090 Wien Austria
- Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., Simpson, I. C., 2007. Final Report for LCM2007 - the new UK Land Cover Map. Countryside Survey Technical Report No 11/07
- Valcarcel, N., Villa, G., Arozarena, A., Garcia-Asensio, L., Caballlero, M., Porcuna, A., Domenech, E. and Peces, J., (2008) SIOSE, a successful test bench towards harmonization and integration of land cover/use information as environmental reference data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37, 1159-1164
- Zimmermann, J., Lydon, K., Packham, I., Smith, G. and Green, S., 2016. The Irish Land-Parcels Identification System (LPIS)–Experiences in ongoing and recent environmental research and land cover mapping. In *Biology and Environment: Proceedings of the Royal Irish Academy* (Vol. 116, No. 1, pp. 53-62). Royal Irish Academy.

Appendices

Appendix 1: Summary definitions for Level 2 classes

Code	Level 2	Short definition
110	Buildings	Intact, constructed, buildings, built-up above ground level, enclosed by walls with a roof or other covering. A structure that is identifiable as having once been a Building but which no longer has a roof, is included as a building. This class is derived directly from PRIME2.
120	Ways	All surfaces mapped as part of the national road and way network in the OSi PRIME2 database.
130	Other Artificial Surfaces	All artificially constructed or synthetic surfaces that are not part of the PRIME2 Building or Way layers.
210	Exposed Rock and Sediments	All exposed consolidated rock in any setting and unconsolidated rock and sediments of any type or size that do not occur on the coastline or as part of an intertidal mudflat.
220	Coastal Sediments	Exposed unconsolidated rock and mineral particles that lie within the coastal zone, in this case defined as a buffer zone of 250 m inland from the edge of the marine water.
230	Mudflats	Large, flat, or gently sloping banks of mud and fine-grained sediments which occur in marine inlets and other low-energy environments at the terrestrial / marine interface.
240	Bare Soil and Disturbed Ground	Surfaces where the top layer of vegetation is broken, disturbed, removed entirely, or worn away and the underlying soil is visible at the time of observation.
250	Burnt Areas	Areas where the vegetation cover has been completely burned away and it is no longer possible to discern what the original vegetation / land cover type once was.
310	Cultivated Land	Land which is recorded in the Land Parcel Information System (LPIS) as being used for the growing of arable crops or horticulture in the reference period of the survey. Also, any other land which is tilled or where arable crops are present in observation of imagery from the same period.
410	Coniferous Forest	Mature forest stands where coniferous tree species make up 75% of more of the forest parcel. The maximum height of the canopy should be greater than 8 m in height.
420	Mixed Forest	Mature forest parcels where neither coniferous nor broadleaved tree species make up 75% of more of the forest parcel. The maximum height of the canopy should be greater than 8 m in height.
430	Transitional Forest	Parcels of plantation (coniferous or mixed as identified through in-situ data) forestry land which has a maximum height of less than 8 m.
440	Broadleaved Forest and Woodland	Mature forest parcels where broadleaved tree species make up 75% of more of the forest parcel. The mean height of the canopy should be greater than 5 m in height.
450	Scrub	Woody shrub vegetation dominated by immature or stunted trees, Gorse, briars and brambles. Vegetation should not exceed 5 m in height.
460	Hedgerows	Narrow linear, generally interconnected lengths of woody shrub vegetation, typically found along field or property boundaries, which have a mean height less than 5 m and a cross-sectional width at the canopy of 10 m or less.
470	Treelines	Linear lengths of mature trees typically found along field or property boundaries, which have a mean height greater than 5 m and a cross-sectional width at the canopy of less than 12 m.

510	Improved Grassland	Agricultural grasslands which are moderately to intensively improved at the time of observation.
520	Amenity Grassland	Any areas of improved grassland which are used for sport, recreation, or other non-agricultural activities. Evidence of such use is determined through in-situ data in the PRIME2 database.
530	Dry Grassland	Low-intensity agricultural and non-agricultural grassland which appears in the imagery to be free draining and dry with minimal or no improvement.
540	Wet Grassland	Low-intensity agricultural and non-agricultural grassland which appears in the imagery to contain an abundance of rushes, sedges and other plants associated with wet grasslands.
550	Saltmarsh	Low-intensity low-lying areas of land dominated by herbaceous vegetation, found adjacent to the intertidal zone or transitional waterbodies and whose substrate is periodically submerged in brackish or saline water.
560	Sand Dunes	Vegetated sections of sand dune systems where marram or other low-intensity / semi-natural grasslands dominate.
570	Swamp	Wetlands which are dominated by stands of herbaceous vegetation which are inundated with standing or slow-moving, brackish, or fresh water throughout the year.
610	Raised Bog	Remaining intact, high bog / dome sections of a raised bog complex, containing *Sphagnum* moss.
620	Blanket Bog	Intact expanses of blanket bog systems.
630	Cutover Bog	Areas of blanket bog or raised bog complexes that show clear evidence of recent or historical cutting / harvesting and are now in a highly degraded state and revegetated to some degree.
640	Bare Peat	Exposed bare peat which can occur as a result of the cutting and harvesting of both raised and blanket bog systems and also by the erosion / weathering of blanket bog in upland locations.
650	Fens	Extremely wet, low-lying areas of land dominated by herbaceous vegetation, mainly grasses, sedges and reeds.
710	Bracken	Areas of open semi-natural vegetation which are dominated by Bracken.
720	Dry Heath	Areas of low (<1.5 m) open vegetation dominated by dwarf shrubs such as Ling (*Calluna vulgaris*), Bell Heather (*Erica cinerea*), Bilberry (*Vaccinium myrtillus*) and low-growing Western Gorse (*Ulex gallii*) in western counties.
730	Wet Heath	Areas of low (<1.5 m) open vegetation with >25% dwarf shrub vegetation, can be dominated by Purple Moor-grass (*Molinia caerulea*) or by Ling (*Calluna vulgaris*) and containing Cross-leaved Heath (*Erica tetralix*).
810	Rivers and Streams	All non-artificial freshwater watercourses which are mapped under the PRIME2 Water layer.
820	Lakes and Ponds	All non-artificial freshwater standing waterbodies that are mapped in the PRIME2 Water layer.
830	Artificial Waterbodies	All artificial watercourses and standing waterbodies that are mapped in the PRIME2 Water layer.
840	Transitional Waterbodies	All intertidal watercourses and standing waterbodies that are mapped in the PRIME2 Water layer.
850	Marine Water	All open sea and marine water that is mapped in the PRIME2 Water layer.

Appendix 2: List of Attributes of Land Cover Data Product

Attributions and definitions for Land Cover Land Use Objects			
Attribute	**Definition**	**Description**	**Source**
Land_Cover_GUID	GUID is a Unique Geographic identifier assigned to each Polygon within Land cover data. Each Land cover GUID could have one or multiple Prime2 GUID – this is 1:1 or 1:Many relationship	A 1: Many relationship in Land Cover means - there could be 1 Prime2 GUID linked to many Land cover Guids/Polygons or 1 Land Cover GUID linked to many Prime2 Guids/Polygons.	Classification.
Level_1_ID	This ID is represented as an Integer showing the ID of the Land Cover Polygon at Level 1 in the classification system.	An integer showing the Land Cover classification at Level 1 in the classification system.	Classification System
Level _1 _Value	This value is represented by Text showing the name of the Land Cover Polygon at Level 1 in the classification system.	This text value displays the Land Cover polygon name at Level 1 in the classification system.	Classification System
Level_1_Capture_Date	Date of when the Polygon was originally captured.	This is the date the data is populated into database.	Final Geodatabase.
Level_1_Change_Code	Where the ID/Name of a Polygon has been reclassified from its original ID/Name to a new ID/Name.	The new ID/Name of the Land Cover Polygon after a change has occurred.	Will be populated when a change occurs in the polygon
Level_1_Change_Date	The date of when the ID/Name of a Polygon was changed to a different ID/Name	Where the name of the polygon has moved from Level 1 to a different name under Level 1	Will be populated when a change occurs in the polygon
Level_1_Capture_Source	Source data that was used to capture the ID/Name of the Polygon e.g. Machine Learning, Prime2, 3rd party data.	How the land cover data was captured. Source data included where applicable.	Linked Tables.
Level_1_Validation_Authority	Organisation who validated the data	Validation partner who carried out the validation of this polygon at Level 1 in the Classification system	Linked Tables.
Level_1 _Validation_Date	Date when the data was validated	The date when the validation partner validated this polygon at Level 1 in the Classification system.	Validated data
Level_2_ID	This ID is represented as an Integer showing the ID of the Land Cover Polygon at Level 2 in the classification system.	An integer showing the Land Cover classification at Level 2 in the classification system.	Classification System
Level_2 _Value	This value is represented by Text showing the name of Land Cover Polygon at Level 2 in the classification system.	This text value displays the Land Cover polygon name at Level 2 in the classification system.	Classification System

Level_2_Capture_Date	Date of when the Polygon was captured.	This is the date the database is populated with data.	Geodatabase.
Level_2_Change_Code	Where the ID/Name of Polygon has been reclassified from its original ID/Name to a new ID/Name.	Discussion need to be held in the future as to how this will be managed in the next iteration.	Will be populated when a change occurs in the polygon
Level_2_Change_Date	The date of when the ID/Name of Polygon was changed to a different ID/Name	Where the name of the polygon has moved from Level 2 to a different name under Level 2	Will be populated when a change occurs in the polygon
Level_2_Capture_Source	Source data that was used to capture the ID/Name of the Polygon e.g. Prime2, 3rd party data, imagery	How the land cover data was captured. Source data included where applicable	Linked Tables.
Level_2_Validation_Authority	Organisation who validated the data	Validation partner who carried out the validation of this polygon at Level 2 in the Classification system	Linked Tables.
Level_2_Validation_Date	Date when the data was validated	The date when the validation partner validated this polygon at Level 2 in the Classification system	Only polygons validated
Capture_Method	The method that was used to capture the data– Machine Learning.	Machine Learning is encapsulating the different datasets that were used during the process– Prime2, 3rd party data, Height data.	Linked Tables.
Image_Resolution	The resolution of the imagery the data was captured from e.g. 25cm 50cm 1m, 10m	The resolution of the Orthophotography used to generate land cover data	Linked Tables.
Capture_Data_Resolution	This relates to MMU – 0.1ha, 0.01ha, 0.05ha, 25/15cm for features from prime2	This could change in the next iteration	Linked Tables.
Image_Name	Name of the image that was used to capture the Objects/Polygons	The name of each Orthophoto used to create each polygon	Ortho Index.
Image_Capture_Date	The date the image was captured	The date the image that is used to create an Orthophoto is taken on	Flight Lines.
Image_Ortho_Series	Which Ortho series was used to capture the data – Series 2, Series 3 etc.	This is an internal OSi production series number. Used to differentiate between capture cycles.	Ortho Index.
Satellite_Image_Name_1	Name of the image that was used	Name of satellite image that was used to create the polygon	Sentinel 2.
Satellite_Image_Date_1	This date is expected to be Spring	A date range to span the date of all the satellite images used to create the polygon	Date range

Satellite_Image_Name_2	Name of the image that was used	Name of satellite image that was used	Sentinel 2
Satellite_Image_Date_2	This date is expected to be Summer	A date range to span the date of all the satellite images used to create the polygon	Date range
Poly_Geom_Capture_Source	This refers to the source of any geometry captured e.g. Post Processing (FME), Prime2.	How the geometry of each polygon was captured. Source data included where applicable.	Geodatabase date, rivers/buildings/ways etc.
Poly_Geom_Capture_Date	The date of the polygon was originally captured	This is the date the data is populated into database.	Prime2 extract.
Poly_Geom_Change_Date	The date the polygon was modified.	The date the geometry was changed	Will be populated when a change occurs in the polygon
Poly_Geom_Validation_Authority	Organisation who validated the data	Validation partner who carried out the validation of the geometry of this polygon.	Linked Tables.
Poly_Geom_Validation_Date	Date when the data was validated	Date when the polygon geometry was validated	Linked Tables
Poly_Geom_Change_Code	Where the code ID/Name of Polygon has been reclassified from its original ID/Name to a new ID/Name e.g. error correction or New Feature.	The date the polygon is changed, a code will signify that a change has occurred.	Will be populated when a change occurs in the polygon

Appendix 3: Minimum Mapping Units for all Level 2 Classes in Final Product

Level 2 Class		MMU
Buildings		As per PRIME2
Ways		0.05 ha
Other Artificial Surfaces		0.01 ha
Exposed Rock and Sediments		0.05 ha
Coastal Sediments		0.05 ha
Mudflats		0.05 ha
Bare Soil and Disturbed Ground		0.05 ha
Burnt Areas		0.05 ha
Cultivated Land		0.05 ha
Coniferous Forest		0.05 ha
Mixed Forest		0.05 ha
Transitional Forest		0.05 ha
Broadleaved Forest and Woodland		0.05 ha
Scrub		0.01 ha
Hedgerows		50 m length
Treelines		50 m length
Improved Grassland		0.05 ha
Amenity Grassland		0.01 ha
Dry Grassland		0.05 ha
Wet Grassland		0.05 ha
Saltmarsh		0.05 ha
Sand Dunes		0.05 ha
Swamp		0.05 ha
Raised Bog		0.05 ha
Blanket Bog		0.05 ha
Cutover Bog		0.05 ha
Bare Peat		0.05 ha
Fens		0.05 ha
Bracken		0.05 ha
Dry Heath		0.05 ha
Wet Heath		0.05 ha
Rivers and Streams		0.05 ha
Lakes and Ponds		0.05 ha
Artificial Waterbodies		0.05 ha
Transitional Waterbodies		0.05 ha
Marine Water		0.05 ha

Appendix 4: Land Cover Level 2 Validation Data Confusion Matrix

Level 2 Class		AG	AW	BP	BSDG	BB	BN	BFW	BG	BA	CS	CF	CL	CB	DG	DH	ERS	FN	HW	IG	LP	MW	MF	MT	OAS	RB	RS	SM	SD	SB	SP	TF	TW	TL	WY	WG	WH	TOTAL
Amenity Grassland	AG	547	0	0	1	0	0	2	3	0	1	1	0	0	10	0	0	0	9	31	0	0	0	0	17	0	1	0	0	8	0	0	0	0	0	5	1	637
Artificial Waterbodies	AW	4	281	0	2	0	0	2	2	0	0	0	0	0	5	0	1	0	0	6	2	0	0	0	15	0	23	2	0	1	2	1	0	1	0	2	0	352
Bare Peat	BP	1	1	301	5	14	0	0	1	0	0	0	1	27	2	12	3	0	0	8	0	0	0	0	6	8	0	0	0	5	0	1	0	1	0	4	4	405
Bare Soil and Disturbed Ground	BSDG	2	0	6	236	3	0	2	2	0	0	2	22	6	7	3	5	5	0	43	2	0	0	0	31	4	1	0	1	2	1	3	1	0	7	12	2	411
Blanket Bog	BB	2	0	2	1	476	4	0	0	1	0	0	0	30	6	21	2	0	0	5	0	0	0	0	0	4	0	0	0	4	0	4	0	0	2	31	56	651
Bracken	BN	3	0	0	2	5	292	1	0	0	0	1	3	3	17	2	2	1	0	9	0	0	0	0	0	0	0	0	0	28	4	0	0	0	1	22	17	413
Broadleaved Forest and Woodland	BFW	12	0	0	0	0	0	450	0	0	0	11	0	1	1	2	0	0	5	7	0	0	6	0	4	0	0	0	0	22	0	12	0	77	0	5	2	617
Buildings	BG	0	0	0	1	0	0	0	423	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	1	0	429
Burnt Areas	BA	1	0	7	0	10	2	0	0	97	1	0	0	0	3	51	5	0	0	2	0	0	0	0	1	0	0	0	0	4	0	1	0	0	0	9	18	212
Coastal Sediments	CS	1	0	0	0	0	1	0	0	0	326	0	0	0	2	0	16	0	0	0	0	4	0	5	3	0	0	0	3	1	0	0	3	0	0	0	0	365
Coniferous Forest	CF	1	0	0	0	0	0	41	0	0	0	536	3	0	1	0	0	0	2	0	0	0	50	0	2	0	0	0	0	4	0	29	0	13	0	0	0	682
Cultivated Land	CL	1	0	0	11	0	0	1	0	0	2	1	660	1	5	1	1	0	1	82	0	0	0	0	5	0	0	0	0	1	1	0	0	0	0	18	0	792
Cutover Bog	CB	3	2	9	15	11	1	1	3	0	0	2	2	320	0	4	0	2	1	23	0	0	0	0	13	9	0	0	0	13	0	5	0	0	0	37	28	504
Dry Grassland	DG	21	0	0	13	1	0	0	1	0	0	0	10	4	252	4	7	0	1	97	0	0	0	0	2	0	0	0	2	5	0	0	0	1	0	34	12	467
Dry Heath	DH	2	0	6	6	25	3	1	0	2	0	3	0	15	7	411	8	3	1	3	0	0	1	0	4	4	1	1	0	19	3	4	0	1	3	40	43	620
Exposed Rock and Sediments	ERS	1	1	2	19	1	0	0	2	0	15	1	2	1	5	7	381	0	0	2	1	1	0	0	33	0	0	1	0	2	0	0	1	0	4	2	8	493
Fens	FN	2	1	0	1	1	0	0	0	0	0	0	0	8	1	1	2	227	0	0	0	0	0	0	0	1	1	1	0	2	5	0	0	0	0	26	9	289
Hedgerows	HW	3	1	0	2	0	0	5	0	0	0	0	0	0	4	0	0	0	585	9	0	0	0	0	3	0	1	0	0	15	0	0	0	30	0	0	1	659
Improved Grassland	IG	86	1	0	11	1	2	7	0	0	0	0	23	2	85	2	2	3	9	3225	0	0	0	0	9	2	0	0	1	9	3	3	0	6	3	230	6	3731
Lakes and Ponds	LP	1	32	3	2	1	1	4	0	0	0	1	0	0	0	0	1	3	0	8	468	0	0	0	3	0	1	0	0	4	2	0	4	0	0	9	1	549
Marine Water	MW	0	0	0	0	0	1	1	0	0	14	0	0	0	1	1	31	0	1	1	0	177	0	10	3	0	4	2	0	0	0	0	17	0	0	2	0	266
Mixed Forest	MF	3	0	0	0	0	0	53	0	0	0	43	0	0	0	0	0	0	0	7	0	0	315	0	3	0	0	0	0	3	0	11	0	2	0	3	0	443
Mudflats	MT	0	0	0	0	0	0	2	0	0	37	0	0	0	0	0	10	0	0	0	0	14	0	298	0	0	1	11	0	2	0	0	3	0	0	3	0	381
Other Artificial Surfaces	OAS	42	0	0	15	0	0	0	6	0	0	0	4	0	2	0	4	0	0	8	0	0	0	0	574	0	0	0	0	1	0	0	0	1	4	1	1	663
Raised Bog	RB	1	1	3	2	7	0	0	0	0	0	0	1	74	1	1	0	0	0	19	0	0	0	0	3	263	0	0	0	5	0	5	0	0	0	30	4	420
Rivers and Streams	RS	0	7	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	1	0	0	441	1	0	0	0	0	2	0	0	0	0	457
Salt Marsh	SM	0	0	0	0	0	0	0	0	0	0	1	2	0	2	0	2	0	1	0	0	0	0	4	1	0	0	366	4	2	18	0	1	0	1	8	1	414
Sand Dunes	SD	2	0	0	1	0	0	0	0	0	2	0	1	0	17	0	0	0	1	9	0	0	0	0	1	0	0	1	291	4	0	0	0	0	0	4	1	335
Scrub	SB	8	1	1	6	2	2	16	0	0	0	2	0	6	12	8	2	0	18	20	0	0	4	0	9	0	0	0	1	375	2	9	0	10	1	36	3	554
Swamp	SP	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	1	12	0	3	9	0	0	2	1	0	0	9	0	2	304	0	0	0	0	30	2	378
Transitional Forest	TF	8	0	0	2	0	2	41	0	0	1	48	3	2	8	2	2	0	2	26	0	0	17	0	4	0	1	1	0	40	0	572	0	2	0	22	4	810
Transitional Waterbodies	TW	0	0	0	0	0	0	0	0	0	8	1	0	0	1	0	26	0	0	0	0	30	0	8	1	0	0	7	0	1	1	0	139	0	0	0	0	223
Treelines	TL	0	0	0	1	0	0	16	0	0	0	1	0	0	0	0	0	0	38	4	0	0	1	0	0	0	0	0	0	8	0	1	0	398	0	1	0	469
Ways	WY	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	505	0	0	512
Wet Grassland	WG	28	2	0	12	5	1	1	1	0	1	0	5	16	38	11	2	0	2	193	0	0	1	0	12	1	2	1	2	27	0	11	0	5	1	933	26	1340
Wet Heath	WH	2	0	1	7	16	12	0	0	0	1	1	3	11	8	17	6	0	0	12	0	0	0	0	1	0	0	0	0	7	0	4	0	0	0	49	513	671
TOTAL		789	332	341	376	579	324	648	444	100	409	657	746	527	504	563	522	256	677	3865	482	226	395	328	769	296	479	404	305	626	346	676	171	548	532	1609	763	21614

Tailte Éireann
Clárúchán, Luacháil, Suirbhéireacht
Registration, Valuation, Surveying

Tailte Éireann
National Mapping Division
Phoenix Park, Dublin 8
D08 F6E4

PHONE
+353-1-802-5300

www.tailte.ie